

10/081935

(FILE 'HCAPLUS' ENTERED AT 09:29:22 ON 04 DEC 2003)

L29 673 SEA FILE=HCAPLUS ABB=ON PLU=ON COCCIDIOID?
L30 242 SEA FILE=HCAPLUS ABB=ON PLU=ON L29 AND (PROTEIN OR
POLYPEPTIDE OR PROTEIN OR PEPTIDE)
L31 44 SEA FILE=HCAPLUS ABB=ON PLU=ON L30 AND VECTOR
L32 14 SEA FILE=HCAPLUS ABB=ON PLU=ON L31 AND RECOMBIN?

L32 ANSWER 1 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN
ACCESSION NUMBER: 2003:846781 HCAPLUS
TITLE: Cholera vaccine candidate 638: intranasal
immunogenicity and expression of a foreign
antigen from the pulmonary pathogen
Coccidioides immitis
AUTHOR(S): Silva, Anisia J.; Mohan, Archana; Benitez, Jorge
A.
CORPORATE SOURCE: Biochemistry and Immunology, Department of
Microbiology, Morehouse School of Medicine, 720
Westview Dr., SW, Atlanta, GA, 30310-1495, USA
SOURCE: Vaccine (2003), 21(32), 4715-4721
CODEN: VACCDE; ISSN: 0264-410X
PUBLISHER: Elsevier Science Ltd.
DOCUMENT TYPE: Journal
LANGUAGE: English
AB *Vibrio cholerae* strain 638 is a live genetically attenuated
candidate cholera vaccine in which the CTX Φ prophage encoding
cholera toxin has been deleted and hapA, encoding an extracellular
Zn-dependent metalloprotease, was insertionally inactivated. Strain
638 was highly immunogenic when inoculated to adult Swiss mice by
the intranasal route as judged by the induction of a strong serum
vibriocidal antibody response. A side-by-side comparison of strain
638 with its isogenic hapA⁺ precursor (strain 81) in the above model
indicated that inactivation of hapA does not affect immunogenicity.
The spherule-associated antigen 2/proline-rich antigen (Ag2/PRA) of
Coccidioides immitis has been shown to protect mice against
coccidioidomycosis to an extent dependent on the modes of
antigen presentation and challenge with *C. immitis* arthrospores. In
this work, we demonstrate the use of a live genetically attenuated
V. cholerae strain to deliver Ag2/PRA. Ag2/PRA was expressed in 638
as a fusion **protein** with the *Escherichia coli* heat labile
toxin B subunit leader **peptide** using the strong Tac
promoter. The **recombinant** Ag2/PRA was efficiently
expressed, processed and secreted to the periplasmic space.
Intranasal immunizations of adult mice with strain 638 expressing
Ag2/PRA induced serum vibriocidal antibody response to the
vector strain and serum total IgG response to Ag2/PRA.
Strain 638 expressing PRA could be recovered from trachea and lung
up to 20 h after immunization but was effectively cleared 72 h
post-inoculation.

L32 ANSWER 2 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN
ACCESSION NUMBER: 2003:454501 HCAPLUS
DOCUMENT NUMBER: 139:35072
TITLE: **Vectors** comprising nucleotide
sequences for target immunogen, PI31, CIIIA and
antisense HERNA mRNA, and their uses including
use as vaccines
INVENTOR(S): McCreavy, David Thomas; Fraser, William Duncan;
Gallagher, James Anthony

10/081935

PATENT ASSIGNEE(S): University of Liverpool, UK
SOURCE: PCT Int. Appl., 52 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2003048371	A2	20030612	WO 2002-GB5512	20021206
WO 2003048371	A3	20030912		

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.: GB 2001-29338 A 20011207
GB 2002-23829 A 20021012

AB The invention provides **vectors** (such as viral **vectors**, plasmid **vectors** or phagemids) comprising:
(a) a heterologous nucleotide sequence encoding an antigenic **polypeptide** from a pathogen (such as viral, bacterial, parasitic or fungal); (b) a nucleotide sequence encoding a protease inhibitor (such as human PI31); a nucleotide sequence for a constitutive, regulatable, and/or cell/tissue-specific promoter; and (d) a nucleotide sequence encoding an inhibitory RNA mols., specifically an antisense human HERNA oligonucleotides. The invention also provides **vectors** comprising a nucleotide sequence encoding CIITA, a **polypeptide** that stimulates the expression of MHC class II genes. The invention further provides the use of said **vectors** as vaccines in production of an immune response (humoral) to said antigens in an animal, such as human, wherein said vaccination may be against a viral, fungal, bacterial or parasitic disorder. Still further, the invention relates: (a) using said **vectors** in production of antibodies, wherein said antibodies may be of therapeutic and/or of diagnostic use; (b) that said **vectors** may be adapted for expression of humanized or chimeric antibodies; and (c) that said **vectors** may be used to used to immunize animals for production of hybridomas expressing a monoclonal antibody against antigen of interest. Finally, the invention provides the cDNA sequences of mouse CIITA, and human PI31, and partial cDNA sequence of human HERNA helicase. The invention related that the use of said **vectors** containing said sequences can be used to enhance secretion of translated immunogen, and enhance DNA vaccination bias away from an MHC class I event towards MHC class II event. In the examples, the invention presented the construction of two **vectors**, pcDNAFinal and pcDNA6TR-IRES-CIITA, wherein pcDNAFinal contains nucleotide sequences encoding immunogen parathyroid hormone-related **protein** (PTHrP), antisense HERNA mRNA, PI31 and CD4+ T-cell epitope from lymphocytic choriomeningitis virus and wherein

10/081935

pcDNA6TR-IRES-CIITA encodes CIITA. Specifically, the invention related that: (a) antisense HERNA RNA can increase the transcriptional efficiency of **vectors** resulting in greater levels of transgene expression; (b) PI3l can inhibit proteasome digestion of **recombinant** antigen making it more assessable to MHC class II antigens; (c) inclusion of CD4+ T-cell epitope ensured that degraded immunogen-MHC class II complex bound to CD4+ T cells; and (d) inclusion of CIITA **protein** allowed for over-expression of MHC class II antigens.

L32 ANSWER 3 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:637813 HCAPLUS

DOCUMENT NUMBER: 137:180844

TITLE: Yeast Bax-responsive genes for drug target identification in yeast and fungi

INVENTOR(S): Contreras, Roland Henri; Eberhardt, Ines; Luyten, Walter Herman Maria Louis; Reekmans, Rieka Josephina

PATENT ASSIGNEE(S): Janssen Pharmaceutica N.V., Belg.

SOURCE: PCT Int. Appl., 344 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002064766	A2	20020822	WO 2001-EP15398	20011221
WO 2002064766	A3	20030626		
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG			

EP 1346044 A2 20030924 EP 2001-997983 20011221

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL, TR

PRIORITY APPLN. INFO.:

EP 2000-870318 A 20001222

EP 2001-870002 A 20010104

EP 2001-870003 A 20010109

WO 2001-EP15398 W 20011221

AB The cDNAs for Saccharomyces cerevisiae genes responding to BAX gene expression as well as the **proteins** encoded by these cDNAs are disclosed. Addnl., Candida albicans and human homologs of the S. cerevisiae genes/**proteins** are provided. The invention describes the use of nucleic acids and **proteins** which are involved in apoptosis in yeast or fungi for the preparation of medicines for treating diseases associated with yeast or fungi or for the treatment of proliferative disorders or for preventing apoptosis in certain diseases. Methods are provided to identify compds. which selectively modulate the expression or functionality of said

10/081935

proteins in the same or a parallel pathway. Also provided are compds. as well as pharmaceutical compns., medicines and vaccines. The invention also comprises new nucleic acid sequences, probes and primers derived thereof, expression **vectors** and host cells transformed with said **vectors**, **polypeptides** and antibodies raised against said **polypeptides**.

L32 ANSWER 4 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:489694 HCAPLUS

DOCUMENT NUMBER: 137:92359

TITLE: Role of signal sequence in vaccine-induced protection against experimental **coccidioidomycosis**

AUTHOR(S): Jiang, Chengyong; Magee, D. Mitchell; Ivey, F. Douglas; Cox, Rebecca A.

CORPORATE SOURCE: Department of Microbiology, University of Texas Health Science Center at San Antonio, San Antonio, TX, 78229, USA

SOURCE: Infection and Immunity (2002), 70(7), 3539-3545
CODEN: INFIBR; ISSN: 0019-9567

PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The vaccine efficacy of the gene sequence encoding the signal **peptide** of the antigen known as antigen 2 or proline-rich antigen (Ag2/PRA), an immunodominant antigen present in the cell wall of the fungal pathogen **Coccidioides immitis**, was investigated in a murine model of **coccidioidomycosis**. Expression plasmids for Ag2/PRA(1-18) DNA (signal sequence), Ag2/PRA(19-194) DNA (lacking the signal sequence), and Ag2/PRA(1-194) DNA (full length) were inserted in the pVR1012 **vector**, and the constructs were used to vaccinate the highly susceptible BALB/c mouse strain. Immunization with the signal gene sequence significantly reduced the fungal burden in the lungs and spleens of mice 12 days after i.p. challenge with a LD of 2500 C. immitis arthroconidia, to a level comparable to the protection induced in mice immunized with the full-length Ag2/PRA(1-194) DNA. The Ag2/PRA(19-194) gene protected mice but to a significantly lower level than the signal sequence or the full-length Ag2 gene. The immunizing capacity of Ag2/PRA(1-18) was not attributable to a nonspecific immunostimulatory effect of DNA, as evidenced by the fact that mice immunized with a frameshift mutation of Ag2/PRA(1-18) were not protected against challenge. Furthermore, a synthetic **peptide** corresponding to the translated sequence of Ag2/PRA(1-18) DNA protected mice, albeit at a lower level than the Ag2/PRA(1-18) DNA vaccine. The protection induced with the signal gene vaccine correlated with the production of gamma interferon when splenocytes from Ag2/PRA(1-18)-immunized mice were stimulated with **recombinant** full-length Ag2 and was not associated with the production of anti-**Coccidioides** IgG antibody. This is the first study to establish that a signal **peptide** sequence alone, administered as a gene vaccine or synthetic **peptide**, can induce protective immunity against a microbial pathogen.

REFERENCE COUNT: 36 THERE ARE 36 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

10/081935

L32 ANSWER 5 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:366877 HCAPLUS

DOCUMENT NUMBER: 137:137416

TITLE: Complementation of the ODC-Escherichia coli
(EWH319) null mutant by expression of the
Coccidioides immitis ornithine
decarboxylase (CiODC) gene

AUTHOR(S): Guevara-Olvera, L.; Guevara-Gonzalez, R. G.;
Munoz-Sanchez, C. I.; Gonzalez-Chavira, M. M.

CORPORATE SOURCE: Inst. Tecnologico de Celaya, Depto. Ingenieria
Bioquimica, Guanajuato, 38010, Mex.

SOURCE: Informacion Tecnologica (2002), 13(2), 177-181
CODEN: ITECFG; ISSN: 0716-8756

PUBLISHER: Centro de Informacion Tecnologica

DOCUMENT TYPE: Journal

LANGUAGE: Spanish

AB The aim of this study was to determine the function of the ornithine
decarboxylase (CiODC) gene from **Coccidioides immitis**, a
human respiratory fungal pathogen. A fragment of complementary DNA
(cDNA) coding for CiODC **protein** was cloned into the
pET-28b **vector**. The pET-CiODC **recombinant**
plasmid was used to transform the EWH319 strain of Escherichia coli,
which lacks the ODC gene and all others which synthesize polyamines
and shows reduced growth due to this deficiency. Introduction of
the transformant plasmid restored the bacterial growth in minimal M9
medium. The complemented strain EWH319.pET-CiODC showed ODC
activity which was inhibited by 1,4-diamino-2-butanone (DAB) a
competitive inhibitor which is specific for eukaryote ODCs. Based
on the present results it was concluded that expression of the CiODC
gene from **Coccidioides immitis** conferred a fundamental
function for the growth of Escherichia coli.

REFERENCE COUNT: 20 THERE ARE 20 CITED REFERENCES AVAILABLE
FOR THIS RECORD. ALL CITATIONS AVAILABLE
IN THE RE FORMAT

L32 ANSWER 6 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:107512 HCAPLUS

DOCUMENT NUMBER: 136:162279

TITLE: Production of human monoclonal antibodies in
human B-lymphocyte hybridomas expressing an
ectopic telomerase gene

INVENTOR(S): Dessain, Scott K.; Goldsby, Richard A.

PATENT ASSIGNEE(S): Whitehead Institute for Biomedical Research, USA

SOURCE: PCT Int. Appl., 74 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
-----	----	-----	-----	-----
WO 2002010352	A2	20020207	WO 2001-US24591	20010801
WO 2002010352	A3	20030227		
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,			

10/081935

NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR,
TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ,
MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH,
CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE,
TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN,
TD, TG

US 2002045219 A1 20020418 US 2001-759984 20010112
PRIORITY APPLN. INFO.: US 2000-222473P P 20000802
US 2001-759984 A1 20010112

AB The present invention relates to a method of making human monoclonal antibodies through the use of novel hybrid cells. In the invention, hybrid cells are created by combining three elements: a fusion partner cell, a fusion cell (in particular a human B-lymphocyte), and an ectopic telomerase gene. Mammalian cell lines that ectopically express telomerase and methods of using such cell lines in producing novel hybrid cells (hybridomas) that produce human monoclonal antibodies; human monoclonal antibodies produced by such novel hybridomas and DNA constructs useful for producing mammalian cell lines that ectopically express telomerase are described. The expression of an ectopic telomerase gene in hybrid cells formed from primary human B-lymphocytes and fusion partner cells (other human non-B lineage cells) improves their growth rate, level of Ig expression, stability of Ig expression, and the ability to be cloned by limiting dilution. A murine myeloma cell line that ectopically expressed human telomerase was created, as well as murine/human cell hybrids.

L32 ANSWER 7 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:935789 HCAPLUS

DOCUMENT NUMBER: 136:65197

TITLE: Sequences of antigenic **polypeptides** of
staphylococcus aureus and their uses in against
bacterial infection

INVENTOR(S): Foster, Simon; McDowell, Philip; Brummell,
Kirsty; Clarke, Simon

PATENT ASSIGNEE(S): University of Sheffield, UK; Biosynexus Inc.

SOURCE: PCT Int. Appl., 85 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001098499	A1	20011227	WO 2001-GB2685	20010620
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			

10/081935

EP 1292681 A1 20030319 EP 2001-940746 20010620
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC,
PT, IE, SI, LT, LV, FI, RO, MK, CY, AL, TR
BR 2001011823 A 20030610 BR 2001-11823 20010620
NO 2002005838 A 20030218 NO 2002-5838 20021205
US 2003186275 A1 20031002 US 2003-311879 20030318
PRIORITY APPLN. INFO.: GB 2000-14907 A 20000620
 WO 2001-GB2685 W 20010620

AB The invention discloses methods for the identification of antigenic **proteins** expressed by pathogenic microbes, vaccines comprising the **proteins**, **recombinant** methods to manufacture the **proteins** and therapeutic antibodies directed to the **proteins**. In particular, the invention discloses amino acid sequences of staphylococcus aureus antigenic **proteins**, the DNA sequences encoding **polypeptides** and genomic DNA library of staphylococcus aureus. The invention also provides expression **vectors** encoding antigenic **peptides**, methods for the production of the **proteins**, antibodies to the **proteins** as well as methods of preparing the antibodies. The invention further provides vaccine comprising the antigenic **proteins**, pharmaceutical carrier, and adjuvant as well as methods of immunizing animals or humans.

REFERENCE COUNT: 2 THERE ARE 2 CITED REFERENCES AVAILABLE FOR
 THIS RECORD. ALL CITATIONS AVAILABLE IN
 THE RE FORMAT

L32 ANSWER 8 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:309883 HCAPLUS

DOCUMENT NUMBER: 135:91197

TITLE: **Recombinant** urease and urease DNA of
 Coccidioides immitis elicit an
 immunoprotective response against
 coccidioidomycosis in mice

AUTHOR(S): Li, Kun; Yu, Jieh-Juen; Hung, Chiung-Yu;
 Lehmann, Paul F.; Cole, Garry T.

CORPORATE SOURCE: Department of Microbiology and Immunology,
 Medical College of Ohio, Toledo, OH, 43614-5806,
 USA

SOURCE: Infection and Immunity (2001), 69(5), 2878-2887
 CODEN: INFIBR; ISSN: 0019-9567

PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB **Coccidioides immitis** antigens which stimulate a T helper cell 1 (Th1) pathway of host immune response are considered to be essential components of a vaccine against **coccidioidomycosis**. **Recombinant** urease (rURE) and **recombinant** heat shock **protein** 60 (rHSP60) of *C. immitis* were expressed in *Escherichia coli* and tested as vaccine candidates in BALB/c mice. A synthetic oligodeoxynucleotide which contained unmethylated CpG dinucleotides and was previously shown to enhance a murine Th1 response was used as an immunoadjuvant. T cells isolated from the spleens and lymph nodes of the rURE- and rHSP60-immune mice showed in vitro proliferative responses to the resp. **recombinant protein**, but only those T lymphocytes from rURE-immunized mice revealed markedly elevated levels of expression of selected Th1-type cytokine genes. BALB/c mice immunized s.c. with rURE and subsequently challenged by the i.p.

10/081935

route with a lethal inoculum of *C. immitis* arthroconidia demonstrated a significant reduction in the level of *C. immitis* infection compared to control animals. RHSP60 was much less effective as a protective antigen. Evaluation of cytokine gene expression in lung tissue and levels of **recombinant** urease-specific Igs (IgG1 vs. IgG2a) in murine sera at 12 days after challenge provided addnl. evidence that immunization with rURE stimulated a Th1 response to the pathogen. Urease was further evaluated by expression of the URE gene in a mammalian plasmid **vector** (pSecTag2A.URE) which was used to immunize mice by the intradermal route. In this case, 82% of the **vector** construct-immunized animals survived more than 40 days after i.p. infection, compared to only 10% of the mice immunized with the **vector** alone. In addition, 87% of the pSecTag2A.URE-immunized survivors had sterile lungs and spleens. These data support the need for further evaluation of the *C. immitis* urease as a candidate vaccine against **coccidioidomycosis**.

REFERENCE COUNT: 44 THERE ARE 44 CITED REFERENCES AVAILABLE
FOR THIS RECORD. ALL CITATIONS AVAILABLE
IN THE RE FORMAT

L32 ANSWER 9 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1999:83280 HCAPLUS

DOCUMENT NUMBER: 130:266091

TITLE: Genetic vaccination against **Coccidioides**
immitis: comparison of vaccine efficacy of
recombinant antigen 2 and antigen 2 cDNA

AUTHOR(S): Jiang, Chengyong; Magee, D. Mitchell; Quitugua,
Teresa N.; Cox, Rebecca A.

CORPORATE SOURCE: Department of Clinical Investigation, Texas
Center for Infectious Disease, San Antonio, TX,
78223, USA

SOURCE: Infection and Immunity (1999), 67(2), 630-635
CODEN: INFIBR; ISSN: 0019-9567

PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Previous studies from our laboratory established that C-ASWS, an alkali-soluble, water-soluble extract from cell walls of **Coccidioides immitis**, protects mice against lethal challenge with this fungus. The C-ASWS extract contains a glycosylated **protein**, designated antigen 2 (Ag2), and a polysaccharide antigen. We recently cloned Ag2 cDNA and showed that the **recombinant** fusion **protein** elicited strong delayed-type hypersensitivity responses in immunized mice. This investigation was undertaken to determine if the **recombinant** Ag2 **protein**, expressed as an Ag2-glutathione S-transferase (GST) fusion **protein**, or Ag2 cDNA would protect mice against lethal challenge with *C. immitis*. The **recombinant** Ag2-GST **protein** protected BALB/c mice against i.p. challenge with 250 arthroconidia, as assessed by a decrease in fungal CFU in tissues. The Ag2-GST-immunized mice did not show, however, an increased survival during a 30-day period postinfection. By contrast, immunization of mice with Ag2 cDNA ligated into the pVR1012 plasmid engendered protection against i.p. challenge with 2,500 arthroconidia and against pulmonary challenge with 50 arthroconidia. Vaccine efficacy paralleled the development of delayed-type hypersensitivity responses to *C. immitis* antigen.

10/081935

Whereas mice vaccinated with the **recombinant** Ag2-GST **protein** did not mount footpad hypersensitivity to C-ASWS or the **recombinant** Ag2-GST **protein**, mice vaccinated with the pVR1012-Ag2 construct mounted a strong footpad hypersensitivity and their spleen cells secreted gamma interferon upon in vitro stimulation with the Ag2-containing C-ASWS extract This is the first investigation to show that genetic immunization can protect against lethal challenge with C. immitis.

REFERENCE COUNT: 50 THERE ARE 50 CITED REFERENCES AVAILABLE
FOR THIS RECORD. ALL CITATIONS AVAILABLE
IN THE RE FORMAT

L32 ANSWER 10 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1999:18776 HCAPLUS

DOCUMENT NUMBER: 130:236203

TITLE: Detecting serum antibodies to a purified
recombinant proline-rich antigen of
Coccidioides immitis in patients with
coccidioidomycosis

AUTHOR(S): Orsborn, Kris I.; Galgiani, John N.

CORPORATE SOURCE: Research and Medical Services, Veterans Affairs
Medical Center, University of Arizona, Tucson,
AZ, USA

SOURCE: Clinical Infectious Diseases (1998), 27(6),
1475-1478

CODEN: CIDIEL; ISSN: 1058-4838

PUBLISHER: University of Chicago Press

DOCUMENT TYPE: Journal

LANGUAGE: English

AB In previous work, antibodies in serum samples from patients with **coccidioidomycosis** were found to react with a proline-rich antigen (PRA) isolated from spherules of **Coccidioides immitis**, and the gene encoding this antigen was cloned. We expressed and purified **recombinant** PRA (rPRA) by removing the majority of amino acids contributed by the **vector** from the fusion **protein**. Purified rPRA reacted with serum IgG antibodies in 37 of 42 patients with culture-proven progressive pulmonary or extrapulmonary **coccidioidal** disease; specific antibodies in dilns. ranging from 1:40 to 1:102,400 were demonstrated (sensitivity, 88%). In contrast, for >95% of patients without **coccidioidomycosis** reactivity of <1:40 was demonstrated (specificity, 97%). Of 18 patients with primary self-limited **coccidioidomycosis**, none had detectable antibodies in serum samples collected up to 141 days after illness began. The association of antibodies to rPRA with progressive infection may have prognostic value.

REFERENCE COUNT: 22 THERE ARE 22 CITED REFERENCES AVAILABLE
FOR THIS RECORD. ALL CITATIONS AVAILABLE
IN THE RE FORMAT

L32 ANSWER 11 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1996:391996 HCAPLUS

DOCUMENT NUMBER: 125:55750

TITLE: Molecular cloning and characterization of
Coccidioides immitis antigen 2 cDNA

AUTHOR(S): Zhu, Yufan; Yang, Chunmu; Magee, D. Mitchell;
Cox, Rebecca A.

CORPORATE SOURCE: Dep. Clinical Investigation, Texas Center

10/081935

SOURCE: Infectious Disease, San Antonio, TX, 78223, USA
Infection and Immunity (1996), 64(7), 2695-2699
CODEN: INFIBR; ISSN: 0019-9567
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Previous expts. have provided evidence that **Coccidioides** immitis antigen 2 (Ag2) is a major T-cell-reactive component of mycelia and spherule cell walls. Here the authors report the identification and cloning of the cDNA that encodes Ag2 from a lambda ZAP cDNA expression library constructed from spherule-derived RNA. DNA sequence anal. established that the 1,255-bp clone contains a 174-bp 5' untranslated region, a 582-bp open reading frame which encodes for a **protein** consisting of 194 amino acids, and a 375-bp 3' untranslated region, including a poly(A) tail. The **recombinant** Ag2 **protein** has a predicted mol. mass of 19.5 kDa and contains an 18-amino-acid N terminus which has been tentatively identified as a signal **peptide**. The Ag2 cDNA was ligated into the pGEX-4T-3 **vector** and expressed in Escherichia coli TG-1 cells as a glutathione S-transferase fusion **protein**. The **recombinant** fusion **protein** showed reactivity with sera from patients with **coccidioidomycosis** and elicited delayed-type footpad hypersensitivity responses in **Coccidioides**-immune mice. These results suggest that the Ag2 cDNA can be used for the large-scale production of this immunol. important **protein**.

L32 ANSWER 12 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1996:327965 HCAPLUS

DOCUMENT NUMBER: 125:28509

TITLE: Molecular cloning and characterization of the **Coccidioides** immitis complement fixation/chitinase antigen

AUTHOR(S): Yang, Chunmu; Zhu, Yufan; Magee, D. Mitchell; Cox, Rebecca A.

CORPORATE SOURCE: Department Clinical Investigation, Texas Center Infectious disease, San Antonio, TX, 78223, USA

SOURCE: Infection and Immunity (1996), 64(6), 1992-1997
CODEN: INFIBR; ISSN: 0019-9567

PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Detection of anti-**Coccidioides** complement-fixing (CF) antibody is a valuable diagnostic and prognostic aid in **coccidioidomycosis**. The CF antibody response is directed against a heat-labile antigen that has chitinase activity, hereafter referred to as the CF/chitinase **protein**. To identify and clone this immunoreactive enzyme, the authors constructed a **Coccidioides** immitis cDNA lambda ZAP expression library from spherule RNA and detected fusion **peptides** expressing CF epitopes by immunoscreening. A cDNA clone consisting of 1623 bp was identified, sequenced, and found to contain a single open reading frame that encodes a **protein** of 47 kDa with 427 amino acids. Deduced amino acid sequence analyses showed that the cloned CF/chitinase cDNA contains a 35-amino-acid region, beginning at Ser-18 and ending at Arg-52, which has 92% homol. with the reported N-terminal amino acid sequence of authentic CF/chitinase

10/081935

protein. The first 17 amino acids in the deduced sequence of the cloned cDNA are not present on the mature CF/chitinase **protein**, suggesting that it may be a signal **peptide**

. Expression of the CF/chitinase cDNA insert by using the pGEX-4T-3 **vector** yields a fusion **peptide** that bears CF-specific epitopes and shows chitinase activity. The CF/chitinase clone will enable large-scale production of the **recombinant** CF antigen for use in immunoassays and facilitate studies on the role of chitinase in the morphogenesis of *C. immitis*.

L32 ANSWER 13 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1995:832234 HCAPLUS

DOCUMENT NUMBER: 124:2086

TITLE: Molecular and biochemical characterization of a **Coccidioides immitis**-specific antigen

AUTHOR(S): Pan, Shuchong; Cole, Garry T.

CORPORATE SOURCE: Department of Microbiology, Medical College of Ohio, Toledo, OH, 43699-008, USA

SOURCE: Infection and Immunity (1995), 63(10), 3994-4002
CODEN: INFIBR; ISSN: 0019-9567

PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Results of earlier investigations have indicated that the saprobic phase of **Coccidioides immitis** produces a heat-stable, 19-kDa antigen with serine proteinase activity which has been suggested to be specific for this pathogenic fungus. In the present study the authors have determined the N-terminal and partial internal amino acid sequences of the purified, 19-kDa antigen, cloned the gene which encodes this **polypeptide**, and confirmed that the secreted proteinase is a **Coccidioides**-specific antigen (CS-Ag). Both the genomic and cDNA sequences are reported and reveal that the *csa* gene which encodes this antigen has no introns. A 543-bp open reading frame encodes a 181-amino-acid-containing **protein** with a predicted mol. mass of 19.8 kDa and an isoelec. point of 8.3. The *csa* gene was localized on chromosome I of three representative *C. immitis* clin. isolates on the basis of Southern hybridizations. Expression of the *csa* gene in *Escherichia coli* using the pET21a plasmid **vector** yielded a **recombinant protein** that was recognized in immunoblot assays by antibody raised to the purified 19-kDa CS-Ag. Secretion of the native antigen is suggested to occur by cleavage of a putative 23-residue signal **peptide**. The native CS-Ag showed a low degree of glycosylation. Anal. of the carbohydrate composition of the CS-Ag revealed xylose, mannose, galactose, and glucose. However, the purified antigen showed no affinity for Con A. A PCR method with specificity and high sensitivity for detection of *C. immitis* genomic DNA, using a pair of synthetic oligonucleotide primers whose sequences were based on that of the *csa* gene, was developed. A 520-bp product was amplified only when *C. immitis* genomic DNA was used as the template. The lower limits of DNA detection using this PCR method were 1 pg of *C. immitis* genomic DNA by ethidium bromide staining and 100 fg after Southern hybridization. The *csa* gene-based PCR method for detection of *C. immitis* DNA is useful for culture identification and may have clin. applications for the diagnosis of **coccidioidal** infections.

L32 ANSWER 14 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

10/081935

ACCESSION NUMBER: 1994:267638 HCAPLUS
DOCUMENT NUMBER: 120:267638
TITLE: Identification of antigens of
Coccidioides immitis which stimulated
immune T lymphocytes
AUTHOR(S): Cole, Garry T.; Kirkland, Theo N.
CORPORATE SOURCE: Dep. Bot., Univ. Texas, Austin, TX, 78713, USA
SOURCE: Archives of Medical Research (1993), 24(3),
281-91
CODEN: AEDEER; ISSN: 0188-4409
DOCUMENT TYPE: Journal
LANGUAGE: English
AB T-cell mediated immune response to **coccidioidomycosis** has
been shown to be the principal mechanism of resistance to this
respiratory fungal disease in exptl. animals. In this study, a
Coccidioides immitis antigen-specific murine T-cell line was
used to identify macromols. capable of eliciting an immune mouse
T-cell proliferative response. The murine T-cell line was selected
on the basis of its strong pos. response to a soluble conidial wall
fraction (SCWF), which had previously been shown to be reactive in
humoral and cellular immunoassays. An antigen-specific T-cell line
rather than T-cell clones was used to identify multiple antigens.
The T-cell immunoblot method was employed first to identify
immunoreactive subfractions of the SCWF, and then to identify T-cell
fusion **proteins** (FPs) obtained from a cDNA expression
library constructed in λ gt11. The library was screened with
anti-SCWF. The nucleotide sequence of a 0.2-kb cDNA insert encoding
a FP which elicited vigorous T-cell response was determined. A construct
of this insert was subcloned into the pET expression **vector**
system and 6.5-kDa **recombinant protein** (RP)
expressed in *Escherichia coli* was isolated. The RP and FP were
shown to be homologous on the basis of identify of their amino acid
sequences. Antibody raised in guinea pigs against the RP recognized
a 59-kDa native **protein** of the mycelial culture filtrate
produced by 3 sep. strains of *C. immitis*, and reacted with the cell
wall of arthroconidia as detected by immunofluorescence microscopy.
In this study, a potentially important T-cell stimulating antigen of
C. immitis was identified and partly characterized.

(FILE 'MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH,
JICST-EPLUS, JAPIO' ENTERED AT 09:34:38 ON 04 DEC 2003)

L33 37 S L32
L34 18 DUP REM L33 (19 DUPLICATES REMOVED)

L34 ANSWER 1 OF 18 WPIDS COPYRIGHT 2003 THOMSON DERWENT on STN
ACCESSION NUMBER: 2003-829776 [77] WPIDS
CROSS REFERENCE: 1999-551417 [46]; 2002-606631 [65]
DOC. NO. NON-CPI: N2003-662909
DOC. NO. CPI: C2003-233753
TITLE: Novel chitin-binding, chitinase-inactive
polypeptide comprising a chitin-binding
fragment, useful for treating fungal infections.
DERWENT CLASS: B04 D16 S03
INVENTOR(S): GRAY, P W; TJOELKER, L W
PATENT ASSIGNEE(S): (GRAY-I) GRAY P W; (TJOE-I) TJOELKER L W
COUNTRY COUNT: 1
PATENT INFORMATION:

10/081935

PATENT NO	KIND	DATE	WEEK	LA	PG
US 2003143216	A1	20030731	(200377)*		33

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
US 2003143216	A1 CIP of	US 1998-39198	19980312
	Div ex	US 1999-267574	19990312
		US 2002-161547	20020603

FILING DETAILS:

PATENT NO	KIND	PATENT NO
US 2003143216	A1 CIP of	US 6200951
	Div ex	US 6399571

PRIORITY APPLN. INFO: US 1999-267574 19990312; US 1998-39198
19980312; US 2002-161547 20020603

AN 2003-829776 [77] WPIDS
CR 1999-551417 [46]; 2002-606631 [65]
AB US2003143216 A UPAB: 20031128

NOVELTY - A chitin-binding, chitinase-inactive **polypeptide** (I) comprising a chitin-binding fragment of 54 C-terminal amino acids of a fully defined human chitinase sequence (S1) of 466 amino acids as given in the specification, is new. (I) is chosen from **polypeptides** having sequence of amino acid residues 347-397 to 445 of (S1).

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for:

- (1) a fusion **protein** (II) comprising (I) fused to a heterologous **polypeptide**;
- (2) a composition (III) comprising (I) and a diluent;
- (3) a composition (IV) comprising (I) conjugated to an antifungal agent;
- (4) a composition (V) comprising (I) conjugated to a detectable label;
- (5) a kit (VI) for diagnosing the presence of chitin in a sample comprising the composition;
- (6) a purified, isolated polynucleotide (VII) encoding (I);
- (7) a **vector** (VIII) comprising (VII);
- (8) a host cell (IX) transformed or transfected with (VII) in a manner allowing the expression of (I) encoded by (VII) in (IX);
- (9) a **polypeptide** produced by culturing (IX); and
- (10) a monoclonal antibody (X) that specifically binds to an epitope within the 54C-terminal amino acid of (S1).

ACTIVITY - Fungicide.

Anti-fungal activity of **recombinant** chitinase in vivo in mice was determined as follows. Female Balb/c mice, 6-8 weeks old, were administered 0.5, 5.0, 50 mg/kg **recombinant** human chitinase by intravenous injection in the tail vein. The mice was infected with the fungi, varying doses of test drug was administered to the animals and their survival is measured over time. Specifically, acute systemic candidiasis was achieved in mice by intraperitoneal or intravenous challenge of 10 multiply 10⁶ colony forming units (CFU) candida albicans. The **recombinant**

10/081935

chitinase was administered before or at 1 to 5 hours after challenge, and the number of survivors was determined after five days. In addition, the mice can be sacrificed and fungal load can be determined in specific organs. Alternatively, the mice were challenged with lower doses of fungi, e.g., *Aspergillus* (8-10 multiply 10⁶ CFU) or *Candida* (1 multiply 10⁶ CFU) in which case survival can be measured at more distant time points, e.g. 45 days. The long term fungicidal/fungistatic activity of a chitinase may be evaluated by continuing therapy for a week or more, e.g. 11 days, and following the animals over several weeks, e.g. 18 days to one month. The result showed that effective anti-fungal agent (**recombinant** chitinase) enhanced the long term survival of animals and reduced fungal load in blood and organs.

MECHANISM OF ACTION - Lysis of fungal cell wall and membrane.

USE - (III) or (IV) is useful for treating (M1) fungal infection which involves administering (III) or (IV) to a subject suffering from fungal infection. (M1) further comprises administering a non-chitinase antifungal agent. (V) is useful for detecting the presence of chitin in a sample which involves contacting the sample with (V) and determining the amount of labeled **polypeptide** bound to chitin. (claimed).

(III), (IV) or (V) is useful for treating fungal infections such as candidiasis, aspergillosis, **coccidioidomycosis**, blastomycosis, paracoccidioidomycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis, and the dermatophytoses can manifest as acute or chronic disease. (VII) is useful in hybridization assays to detect the capacity of cells to synthesize chitinase and for diagnostic methods useful for identifying a genetic alteration(s) in the chitinase locus that underlies a disease state or states. (X) is useful for detecting or quantifying the presence of chitin-binding domain, and for detecting or quantifying the presence of yeast or fungi, e.g., by adding a chitin-binding domain which binds to the yeast or fungi. (X) is also useful for detecting chitin-binding domain in human blood samples which indicates a disease state involving chitinase, such as Gaucher's disease.

Dwg.0/0

L34 ANSWER 2 OF 18 MEDLINE on STN DUPLICATE 1
ACCESSION NUMBER: 2003510266 IN-PROCESS
DOCUMENT NUMBER: 22949199 PubMed ID: 14585681
TITLE: Cholera vaccine candidate 638: intranasal
immunogenicity and expression of a foreign antigen
from the pulmonary pathogen **Coccidioides**
immitis.
AUTHOR: Silva Anisia J; Mohan Archana; Benitez Jorge A
CORPORATE SOURCE: Department of Microbiology, Biochemistry and
Immunology, Morehouse School of Medicine, 720
Westview Dr., SW 30310-1495, Atlanta, GA, USA.
SOURCE: VACCINE, (2003 Dec 1) 21 (32) 4715-21.
Journal code: 8406899. ISSN: 0264-410X.
PUB. COUNTRY: Netherlands
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: IN-PROCESS; NONINDEXED; Priority Journals
ENTRY DATE: Entered STN: 20031031
Last Updated on STN: 20031031
AB *Vibrio cholerae* strain 638 is a live genetically attenuated

10/081935

candidate cholera vaccine in which the CTXPhi prophage encoding cholera toxin has been deleted and hapA, encoding an extracellular Zn-dependent metalloprotease, was insertionally inactivated. Strain 638 was highly immunogenic when inoculated to adult Swiss mice by the intranasal route as judged by the induction of a strong serum vibriocidal antibody response. A side-by-side comparison of strain 638 with its isogenic hapA(+) precursor (strain 81) in the above model indicated that inactivation of hapA does not affect immunogenicity. The spherule-associated antigen 2/proline-rich antigen (Ag2/PRA) of *Coccidioides immitis* has been shown to protect mice against *coccidioidomycosis* to an extent dependent on the modes of antigen presentation and challenge with *C. immitis* arthrospores. In this work, we demonstrate the use of a live genetically attenuated *V. cholerae* strain to deliver Ag2/PRA. Ag2/PRA was expressed in 638 as a fusion **protein** with the *Escherichia coli* heat labile toxin B subunit leader **peptide** using the strong Tac promoter. The **recombinant** Ag2/PRA was efficiently expressed, processed and secreted to the periplasmic space. Intranasal immunizations of adult mice with strain 638 expressing Ag2/PRA induced serum vibriocidal antibody response to the **vector** strain and serum total IgG response to Ag2/PRA. Strain 638 expressing PRA could be recovered from trachea and lung up to 20h after immunization but was effectively cleared 72h post-inoculation.

L34 ANSWER 3 OF 18 WPIDS COPYRIGHT 2003 THOMSON DERWENT on STN
ACCESSION NUMBER: 2003-120579 [11] WPIDS
DOC. NO. NON-CPI: N2003-096057
DOC. NO. CPI: C2003-031163
TITLE: Identifying biologically active agents comprises cloning transfected cells into a cell array, exposing the array to an agent to be tested, and detecting signals generated by a reporter molecule as a result of exposure to the agent.
DERWENT CLASS: B04 D16 S03
INVENTOR(S): ANDREWS, P; DRAPER, J; WALSH, J
PATENT ASSIGNEE(S): (AXOR-N) AXORDIA LTD
COUNTRY COUNT: 100
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG

WO 2002090992	A2	20021114	(200311)*	EN	90
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC					
MW MZ NL OA PT SD SE SL SZ TR TZ UG ZM ZW					
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ					
DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP					
KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ					
NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ					
UA UG US UZ VN YU ZA ZM ZW					

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE

WO 2002090992	A2	WO 2002-GB1946	20020429

10/081935

PRIORITY APPLN. INFO: GB 2001-11004 20010504

AN 2003-120579 [11] WPIDS

AB WO 200290992 A UPAB: 20030214

NOVELTY - Screening (M1) for identifying biologically active agents, comprises:

(i) providing a population of cells stably transfected with a nucleic acid encoding a reporter molecule;

(ii) cloning the transfected cells into a cell array;

(iii) exposing the array to at least one agent to be tested;

and

(iv) detecting a signal generated by the reporter molecule as a result of exposure to the agent.

DETAILED DESCRIPTION - Screening (M1) for identifying biologically active agents, comprises:

(i) providing a population of cells which have been stably transfected with a nucleic acid molecule encoding a reporter molecule;

(ii) cloning the transfected cells into a cell array;

(iii) exposing the array to at least one agent to be tested;

and

(iv) detecting a signal generated by the reporter molecule as a result of exposure to the agent.

INDEPENDENT CLAIMS are also included for the following:

(1) An agent identified by M1;

(2) A cell or a cell array obtained by M1;

(3) Screening (M2) for the isolation of a gene, comprising:

(a) steps (i)-(iv) of M1;

(b) extracting nucleic acid from a cell sample comprising the cell array; and

(c) determining the sequence of at least part of the genomic region into which the nucleic acid encoding the reporter molecule has integrated;

(4) Comparing the biological activity of a reference agent with at least one other agent, comprising:

(a) steps (i) and (ii) of the above method;

(b) preparing a duplicate array;

(c) step (iii) of the above method;

(d) exposing the duplicate array to a reference agent; and

(e) detecting a signal generated by the reporter molecule as a result of exposure to the agent and to the reference agent; and

(5) A **vector** comprising a reporter molecule, a splice acceptor site and an internal ribosome entry site, where the splice acceptor and the internal ribosome entry site are operably linked to facilitate expression of the reporter molecule.

USE - The method is useful in identifying biologically active agents and the genes through which the agents act, in screening potential drugs for their ability to activate certain drug targets in a high-throughput assay, in identifying relationships between signaling pathways and specific signals that could be useful in eventually directing the differentiation of embryonic stem cells, and in toxicology assays by testing for unwanted activation or inhibition of specific signaling pathways. The **vector** is useful in carrying out the above methods (claimed).

Dwg.0/99

L34 ANSWER 4 OF 18 WPIDS COPYRIGHT 2003 THOMSON DERWENT on STN

ACCESSION NUMBER: 2002-566694 [60] WPIDS

DOC. NO. CPI: C2002-160652

10/081935

TITLE: Constructing strains for identifying gene products as effective targets for therapeutic intervention, by inactivating in the strain one allele of a gene and placing other allele of the gene under conditional expression.

DERWENT CLASS: B04 C06 D16

INVENTOR(S): BOONE, C; BUSSEY, H; JIANG, B; OHLSEN, K L; ROEMER, T

PATENT ASSIGNEE(S): (ELIT-N) ELITRA PHARM INC

COUNTRY COUNT: 101

PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2002053728	A2	20020711	(200260)*	EN	167
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZM ZW					
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ UA UG US UZ VN YU ZA ZW					
US 2003180953	A1	20030925	(200364)		
EP 1348027	A2	20031001	(200365)	EN	
R: AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI TR					

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2002053728	A2	WO 2001-US49486	20011226
US 2003180953	A1	US 2000-259128P	20001229
	CIP of	US 2001-792024	20010220
	Provisional	US 2001-314050P	20010822
		US 2001-32585	20011220
EP 1348027	A2	EP 2001-991419	20011226
		WO 2001-US49486	20011226

FILING DETAILS:

PATENT NO	KIND	PATENT NO
EP 1348027	A2 Based on	WO 2002053728

PRIORITY APPLN. INFO: US 2001-314050P 20010822; US 2000-259128P 20001229; US 2001-792024 20010220; US 2001-32585 20011220

AN 2002-566694 [60] WPIDS

AB WO 200253728 A UPAB: 20020919

NOVELTY - Constructing (M1) a strain of diploid fungal cells in which both alleles of a gene are modified, comprising modifying one allele by insertion or replacement by a cassette having an expressible selectable marker, and modifying other allele by **recombination**, of a promoter replacement fragment with a heterologous promoter, so that expression of the second allele is regulated by the promoter, si new.

DETAILED DESCRIPTION - Constructing (M1) a strain of diploid

fungal cells in which both alleles of a gene are modified, comprising constructing a strain of diploid fungal cells in which both alleles of a gene are modified, comprises modifying one allele in diploid fungal cells by **recombination** using a gene disruption cassette comprising a first nucleotide sequence encoding an expressible selectable marker, to provide heterozygous diploid fungal cells in which the first allele is inactivated, and modifying the second allele by **recombination** using a promoter replacement fragment comprising a second nucleotide sequence encoding a heterologous promoter, so that expression of the second allele is regulated by the heterologous promoter, where the gene encodes a **polypeptide** required for proliferation of *Candida albicans*, where the **polypeptide** is chosen from the amino acid sequence (S1) of 932 essential genes listed in the specification. INDEPENDENT CLAIMS are also included for the following:

(1) assembling (M2) a collection of diploid fungal cells each comprising modified alleles of a different gene, by modifying both alleles of the gene by (M1) and repeating the steps several times, where a different gene is modified with each repetition;

(2) a strain (I) of diploid fungal cells comprising modified alleles of a gene, where the first allele is inactivated by a gene disruption cassette comprising a nucleotide sequence encoding an expressible selectable marker, and the expression of the second allele is regulated by a heterologous promoter that is operably linked to the coding region of the second allele;

(3) a collection of diploid fungal strains comprising (I);

(4) a nucleic acid molecule microarray comprising several nucleic acid molecules, each comprising a nucleotide sequence that is hybridizable to a target nucleotide sequence chosen from 932 nucleotide sequences (S2) of the open reading frame (ORF) of essential genes, gene names listed in the specification, no sequences are given;

(5) a purified or isolated nucleic acid molecule (II) comprising a nucleotide sequence encoding a gene product required for proliferation of *C. albicans*, and comprising (S1);

(6) a nucleic acid molecule (III) comprising at least 10-100 consecutive nucleotides of (S2);

(7) a purified or isolated nucleic acid molecule (IV) obtained from an organism other than *C. albicans* or *Saccharomyces cerevisiae* comprising at least 30 % identity to a (S2), as determined using BLASTN version 2.0 with the default parameters;

(8) a **vector** containing the promoter operably linked to (II), (III) or (IV);

(9) a host cell containing the **vector** of (9);

(10) a purified or isolated **polypeptide** comprising a sequence chosen from 61 sequences not given in the specification;

(11) a purified or isolated **polypeptide** (V) obtained from an organism other than *C. albicans* or *S. cerevisiae* comprising at least 30 % similarity to (S1), as determined using FASTA version 3.0t78 with the default parameters;

(12) a fusion **protein** comprising a fragment of a first **polypeptide** fused to a second **polypeptide**, the fragment at least 6 consecutive residues of (S1);

(13) producing (V);

(14) identifying (M3) a compound which modulates the activity of a gene product encoded by a nucleic acid comprising (S2), by contacting the gene product with a compound, and determining if the

compound modulates the activity of the gene product;

(15) eliciting an immune response in an animal, by introducing into the animal a composition comprising an isolated **polypeptide** having 6 consecutive residues of (S1);

(16) a strain of *C. albicans* where a first allele of a gene comprising (S2) is inactive and a second allele is under the control of a heterologous promoter;

(17) identifying a compound or binding partner that binds to a **polypeptide** comprising (S1) or its fragment;

(18) identifying (M4) a compound having the ability to inhibit growth or proliferation of *C. albicans*, by reducing the level or activity of a gene product encoded by (S2) in the *C. albicans* cell relative to a wild type cell, where the reduced level is not lethal to the cell, contacting the cell with a compound and determining if the compound inhibits the growth or proliferation of the cell;

(19) inhibiting (M5) growth or proliferation of *C. albicans* cells, by contacting the cells with the compound that reduce the level of or inhibit the activity of (S1) or (S2);

(20) manufacturing an antimycotic compound, by screening several compounds to identify a compound that reduces the activity or level of a gene product encoded by (S2) and manufacturing the compound identified;

(21) treating (M6) an infection of a subject by *C. albicans*, by administering compound (C) that reduces the activity or level of a gene product encoded by (S2);

(22) an antibody preparation (VI) which binds (V); and

(23) correlating changes in the levels of **proteins** or gene transcripts with the inhibition of growth or proliferation of a diploid fungal cell, by generating a first **protein** expression profile for or transcription profile for a control diploid fungal cell which comprises two wild type alleles of the gene, culturing (I) under conditions where the second allele of the gene is underexpressed, not expressed or overexpressed, and generating a second **protein** expression profile for the cultured cells, and comparing the profiles.

ACTIVITY - Fungicide; Plant Antifungal.

No biological data is given.

MECHANISM OF ACTION - None given.

USE - (M1) is useful for constructing a strain of diploid fungal cells in which both alleles of a gene are modified. The diploid fungal cells having both alleles modified are useful for identifying a gene that is essential to the survival or growth of a fungus, a gene that contributes to the virulence and/or pathogenicity of a fungus, a gene that contributes to the resistance of a diploid fungus to an antifungal agent, an antifungal agent that inhibits the growth of a diploid fungus and for identifying a therapeutic agent for treatment of a mammalian disease. (M3) is useful for identifying a compound which modulates the activity of a gene product, preferably enzymatic activity, carbon compound catabolism, biosynthetic, transporter, transcriptional, translational, signal transduction, DNA replication and cell division activity. (M4) is useful for identifying a compound having the ability to inhibit growth or proliferation of *C. albicans*. (M5) is useful for inhibiting growth or proliferation of *C. albicans* cells. (M6) is useful for treating infection by *C. albicans*. (All claimed).

Dwg.0/6

10/081935

L34 ANSWER 5 OF 18 WPIDS COPYRIGHT 2003 THOMSON DERWENT on STN
ACCESSION NUMBER: 2002-666824 [71] WPIDS
DOC. NO. CPI: C2002-187137
TITLE: Nucleic acid molecules comprising fungal, e.g.
Cochliobolus heterostrophus, genes from a
peptide synthetase gene cluster, useful for
identifying anti-fungal agents for treating fungal
infections such as pneumonia and arthritis.
DERWENT CLASS: C07 D16
INVENTOR(S): LU, S; TURGEON, B G; YODER, O
PATENT ASSIGNEE(S): (CORR) CORNELL RES FOUND INC; (LUSS-I) LU S; (SYGN)
SYNGENTA PARTICIPATIONS AG; (TURG-I) TURGEON B G;
(YODE-I) YODER O
COUNTRY COUNT: 100
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2002042444	A2	20020530	(200271)*	EN	316
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZM ZW					
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZM ZW					
AU 2002036448	A	20020603	(200277)		

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2002042444	A2	WO 2001-US43381	20011121
AU 2002036448	A	AU 2002-36448	20011121

FILING DETAILS:

PATENT NO	KIND	PATENT NO
AU 2002036448	A Based on	WO 2002042444

PRIORITY APPLN. INFO: US 2000-252732P 20001122; US 2000-252649P
20001122

AN 2002-666824 [71] WPIDS
AB WO 200242444 A UPAB: 20030410
NOVELTY - Nucleic acid molecules comprising fungal, e.g.
Cochliobolus heterostrophus, genes from a **peptide**
synthetase gene cluster, e.g. an iron reductase and/or a permease or
a major facilitator superfamily (MFS) transporter, are new.
DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for
the following:
(1) an isolated polynucleotide (N1) comprising a fungal nucleic
acid segment which encodes a **polypeptide** which is
substantially similar to a **polypeptide**, encoded by a
nucleic acid sequence (NA1) comprising an open reading frame
comprising the 2435 (S46), 1836 (S48), or 2073 (S55) base pair (bp)
sequence defined in the specification, or their complements;
(2) an isolated polynucleotide (N2) comprising a fungal nucleic

acid segment which is substantially similar to NAl, or its complements;

(3) an isolated polynucleotide (N3) comprising a fungal nucleic acid segment which hybridizes under stringent hybridization conditions to S46, S48, S55, or its complements;

(4) an isolated **polypeptide** (P1) encoded by N1, N2 or N3;

(5) an expression cassette comprising a promoter operably linked to N1, N2 or N3;

(6) a **recombinant vector** comprising N1, N2 or N3, where the **vector** is capable of being stably transformed into a host cell;

(7) a host cell comprising the expression cassette of (5) or the **vector** of (6);

(8) a method (M1) for identifying an agent having fungicidal or mycocidal activity, comprising contacting a fungus with an agent that binds to or inhibits P1, and identifying an agent having fungicidal or mycocidal activity;

(9) a method (M2) for identifying an inhibitor of a **polypeptide**, comprising contacting a host cell which expresses a **polypeptide** encoded by N1, N2 or N3 with an agent, and identifying an agent that inhibits the activity of the **polypeptide**;

(10) an agent identified by the M1 or M2;

(11) a method of inhibiting the growth or pathogenicity of a fungus, comprising contacting the fungus with the agent of (10) in an amount sufficient to inhibit the growth or pathogenicity of the fungus;

(12) a method for identifying an agent that modulates a **polypeptide** associated with pathogenicity of a fungus, comprising contacting a fungus with an agent that binds P1, and identifying an agent that modulates the pathogenicity of the fungus;

(13) a method for identifying an agent that modulates the pathogenicity of a fungus, comprising contacting a fungus with an agent that inhibits the activity of P1, and identifying an agent that modulates the pathogenicity of the fungus;

(14) a method (M3) of identifying agents that alter the phenotype of a fungal pathogen or mycogen, comprising:

(a) contacting an agent to be tested with one or more cells of a fungal pathogen or mycogen where:

(i) the cells comprise a nucleotide sequence encoding a **polypeptide** that is substantially similar to the 812 (S47), 611 (S49) or 487 (S56) amino acid sequence defined in the specification; or

(ii) the cells have a mutation in N1, N2 or N3, where the mutation results in overexpression or underexpression of the encoded **polypeptide**; and

(b) detecting or determining whether the agent selectively modulates expression or function or metabolic pathways associated with the **polypeptide**, therefore altering a phenotype of the cells relative to cells not contacted with the agent;

(15) an isolated antibody (Ab1) which specifically binds to P1;

(16) a method (M4) to prevent or inhibit infection of an animal or plant by a fungal pathogen, comprising administering to the animal or plant an effective amount of the agent of (10) for a time and under conditions sufficient to inhibit or prevent fungal growth or reproduction;

(17) the sequence of N1, N2 or N3, or the amino acid sequence

10/081935

of P1 stored on a computer readable medium; and

(18) a transformed plant, the genome of which expresses a chimeric DNA molecule which encodes a gene product which confers resistance or tolerance to the plant to a fungal pathogen by inhibiting fungal iron metabolism or siderophore production.

ACTIVITY - Fungicidal; Mycocidal; Antibacterial; Antiarthritic; Dermatological; Antiinflammatory.

No biological data given.

MECHANISM OF ACTION - Antisense therapy.

No biological data given.

USE - The polynucleotides and **polypeptides** are useful for identifying a novel fungicidal or mycocidal mode of action which permits rapid discovery of novel inhibitors of gene products that are useful as fungicides or mycocides.

The agents and antisense DNA are useful as fungicides to suppress the growth of pathogenic fungi. The fungal pathogens include plant pathogens such as *Septoria tritici*, or *Cochliobolus heterostrophus*, or animal pathogens such as *Candida albicans*.

The anti-fungal agents are particularly useful for treating fungal infections of vertebrates, including immunocompromised vertebrates, for e.g. pneumonia, arthritis, miliary disease, bone and joint infection, skin disease, aesophagitis, vaginitis, onychomycosis, and inflammation of urinary tract, kidney, liver, brain, gastrointestinal tract and lung.

Dwg.0/10

L34 ANSWER 6 OF 18 WPIDS COPYRIGHT 2003 THOMSON DERWENT on STN

ACCESSION NUMBER: 2002-739924 [80] WPIDS

DOC. NO. NON-CPI: N2002-582921

DOC. NO. CPI: C2002-209497

TITLE: New **recombinant** spherical outer wall glycoproteins from **Coccidioides immitis**, useful as antigens for serodiagnosis or induction of an immune response.

DERWENT CLASS: B04 C06 C07 D16 S03

INVENTOR(S): COLE, G T; HUNG, C

PATENT ASSIGNEE(S): (MEDI-N) MEDICAL COLLEGE OHIO

COUNTRY COUNT: 1

PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
US 2002106380	A1	20020808	(200280)*		26

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
US 2002106380	A1	Provisional	
		US 2000-202754P	20000508
		US 2001-850677	20010507

PRIORITY APPLN. INFO: US 2000-202754P 20000508; US 2001-850677 20010507

AN 2002-739924 [80] WPIDS

AB US2002106380 A UPAB: 20021212

NOVELTY - A **recombinant** spherical outer wall glycoprotein 58 (SOWgp58) (I) from **Coccidioides immitis**, is new.

Searcher : Shears 308-4994

10/081935

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for:

- (1) antigens (II) for serodiagnosis of **coccidioidomycosis** comprising a **recombinant** spherical outer wall glycoprotein (SOWgp);
- (2) an isolated nucleic acid (III) that hybridizes under highly stringent conditions to a probe of sequences (S1: 900 base pairs), (S3: 1244 base pairs), (S9: 1504 base pairs) or (S11: 1563 base pairs), or their complements;
- (3) purified **polypeptides** (IV) that comprise sequences (S2: 299 amino acids (aa)), (S4: 328 aa), (S5: 15 aa), (S6: 10 aa), (S7: 10 aa), (S8: 14 aa), (S10: 375 aa) or (S12: 422 aa), all (sic) with up to 20 conservative aa substitutions;
- (4) eliciting (M1) a proliferative response from monocytes, in animals that show a positive skin test, by administering Ag; and
- (5) identifying (M2) the SOW gp58, gp66 and/or gp82 genes.

ACTIVITY - Fungicide.

No biological data given.

MECHANISM OF ACTION - Induction of specific humoral and cellular immune responses.

Purified SOWgp58 from the Silveria isolate was tested in an essentially conventional proliferation assay (thymidine incorporation) against peripheral blood mononuclear cells. At a concentration of 50 micro g/ml in the cell culture medium it produced a stimulation index of 16, using cells from subjects who gave a positive skin test for *C. immitis*, but the index was only about 1 when using cells from subjects who gave a negative skin test.

USE - Spherical outer wall glycoproteins (SOW) are antigens useful for:

- (a) serodiagnosis of **coccidioidomycosis** (San Joaquin Valley fever); and
 - (b) for stimulating proliferation of peripheral blood monocytes, for protection against this disease.
- (III) that encode SOW are useful as probes and primers for cloning, and as diagnostic probes.
- Dwg.0/9

L34 ANSWER 7 OF 18 MEDLINE on STN DUPLICATE 2
ACCESSION NUMBER: 2002322692 MEDLINE
DOCUMENT NUMBER: 22060669 PubMed ID: 12065493
TITLE: Role of signal sequence in vaccine-induced protection against experimental **coccidioidomycosis**.
AUTHOR: Jiang Chengyong; Magee D Mitchell; Ivey F Douglas; Cox Rebecca A
CORPORATE SOURCE: Department of Microbiology, University of Texas Health Science Center at San Antonio, San Antonio, Texas 78229, USA.
SOURCE: INFECTION AND IMMUNITY, (2002 Jul) 70 (7) 3539-45. Journal code: 0246127. ISSN: 0019-9567.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200207
ENTRY DATE: Entered STN: 20020615
Last Updated on STN: 20020731
Entered Medline: 20020730

Searcher : Shears 308-4994

10/081935

AB The vaccine efficacy of the gene sequence encoding the signal **peptide** of the antigen known as antigen 2 or proline-rich antigen (Ag2/PRA), an immunodominant antigen present in the cell wall of the fungal pathogen **Coccidioides immitis**, was investigated in a murine model of **coccidioidomycosis**. Expression plasmids for Ag2/PRA(1-18) DNA (signal sequence), Ag2/PRA(19-194) DNA (lacking the signal sequence), and Ag2/PRA(1-194) DNA (full length) were inserted in the pVR1012 **vector**, and the constructs were used to vaccinate the highly susceptible BALB/c mouse strain. Immunization with the signal gene sequence significantly reduced the fungal burden in the lungs and spleens of mice 12 days after intraperitoneal challenge with a lethal dose of 2,500 *C. immitis* arthroconidia, to a level comparable to the protection induced in mice immunized with the full-length Ag2/PRA(1-194) DNA. The Ag2/PRA(19-194) gene protected mice but to a significantly lower level than the signal sequence or the full-length Ag2 gene. The immunizing capacity of Ag2/PRA(1-18) was not attributable to a nonspecific immunostimulatory effect of DNA, as evidenced by the fact that mice immunized with a frameshift mutation of Ag2/PRA(1-18) were not protected against challenge. Furthermore, a synthetic **peptide** corresponding to the translated sequence of Ag2/PRA(1-18) DNA protected mice, albeit at a lower level than the Ag2/PRA(1-18) DNA vaccine. The protection induced with the signal gene vaccine correlated with the production of gamma interferon when splenocytes from Ag2/PRA(1-18)-immunized mice were stimulated with **recombinant** full-length Ag2 and was not associated with the production of anti-**Coccidioides** immunoglobulin G antibody. This is the first study to establish that a signal **peptide** sequence alone, administered as a gene vaccine or synthetic **peptide**, can induce protective immunity against a microbial pathogen.

L34 ANSWER 8 OF 18 WPIDS COPYRIGHT 2003 THOMSON DERWENT on STN
ACCESSION NUMBER: 2001-489080 [53] WPIDS
DOC. NO. NON-CPI: N2001-361848
DOC. NO. CPI: C2001-146921
TITLE: Identifying genes essential to fungal metabolisms
and identifying potential therapeutic agents that
target these genes.
DERWENT CLASS: B04 D16 S03
INVENTOR(S): BOONE, C; BUSSEY, H; JIANG, B; ROEMER, T
PATENT ASSIGNEE(S): (ELIT-N) ELITRA PHARM INC
COUNTRY COUNT: 95
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG

WO 2001060975	A2	20010823	(200153)*	EN	320
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC					
MW MZ NL OA PT SD SE SL SZ TR TZ UG ZW					
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE					
DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG					
KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ					
PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN					
YU ZA ZW					
AU 2001043204	A	20010827	(200176)		
EP 1292668	A2	20030319	(200322)	EN	
R: AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK					

10/081935

NL PT RO SE SI TR
KR 2002097180 A 20021231 (200330)
JP 2003523197 W 20030805 (200353) 370

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2001060975	A2	WO 2001-US5551	20010220
AU 2001043204	A	AU 2001-43204	20010220
EP 1292668	A2	EP 2001-916144	20010220
		WO 2001-US5551	20010220
KR 2002097180	A	KR 2002-710815	20020819
JP 2003523197	W	JP 2001-560347	20010220
		WO 2001-US5551	20010220

FILING DETAILS:

PATENT NO	KIND	PATENT NO
AU 2001043204	A Based on	WO 2001060975
EP 1292668	A2 Based on	WO 2001060975
JP 2003523197	W Based on	WO 2001060975

PRIORITY APPLN. INFO: US 2000-183534P 20000218

AN 2001-489080 [53] WPIDS

AB WO 200160975 A UPAB: 20010919

NOVELTY - Methods for constructing fungal strains useful for identification and validation of gene products as target for therapeutic agents, for identifying and validating gene products as target for therapeutic agents, collections of identified genes and screening assays for the discovery of new drugs, are new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are included for the following:

(1) a method (I) of constructing a strain of diploid fungal cells (DFCs) in which both alleles (Als) of a gene are modified and for assembling a collection of DFCs each of which comprises modified Als of a different gene;

(2) a strain of DFCs comprising modified Als of a gene, wherein the first Al of the gene is inactivated by **recombination** using a gene disruption cassette comprising a nucleotide (Nt) sequence (seq) encoding an expressible selectable marker and the expression of the second Al of the gene is regulated by a heterologous promoter that is operably linked to the coding region of the second Al of the gene;

(3) a collection (III) of diploid fungal strains (II) in which each strain comprises modified Als of a different gene, and all the different genes in the genome of the fungus are modified and represented in the collection;

(4) a NAM microarray (IV) comprising a plurality of nucleic acid molecules (NAMs) (each NAM comprises a Nt seq that is hybridizable to a target Nt seq selected from 62 defined seqs given in the specification, a Nt seq that is hybridizable to the Nt seq of a gene that is either essential to the growth of a DFC or contributes to the virulence and/or pathogenicity of the DFCs against a host organism;

(5) a method (V) for identifying a gene that is essential to the metabolism/life cycle of a fungus comprising:

10/081935

(a) culturing the DFCs (II) under conditions in which the second Al of the gene is substantially under expressed (expd) or not expd; and

(b) determining:

(i) viability of the cells (a loss or reduction of viability as compared to a control indicates that the modified gene is essential to the survival of the fungus);

(ii) growth of the cells (a loss or reduction of growth of the cells as compared to a control indicates that the modified gene is essential to the growth of the fungus);

(iii) growth of the cells (a loss or reduction of growth of the cells as compared to a control indicates that the modified gene is essential to the growth of the fungus);

(iv) viability of the cells (an increase in viability as compared to a control indicates that the modified gene contributes to the resistance of the diploid fungus to the antifungal agent);

(6) a method (VI) for identifying an antifungal agent that inhibits the growth of a diploid fungus;

(7) a method (VII) for identifying a therapeutic agent for treatment of a mammalian disease;

(8) a method (VIII) for correlating changes in the levels of **proteins** with the inhibition of growth or proliferation of a DFC, comprising:

(a) generating a first **protein** expression profile for a control DFC which comprises two wild type Als of the gene;

(b) culturing DFCs (II) under conditions wherein the second Al of the gene is substantially under expd, not expd or over expd, and generating a second **protein** expression profile for the cultured cells; and

(c) comparing the first **protein** expression profile with the second **protein** expression profile to identify changes in the levels of **proteins**;

(9) a method (IX) for correlating changes in the levels of gene transcripts with the inhibition of growth or proliferation of a DFC, comprising:

(a) generating a transcription profile for a control DFC which comprises two wild type Als of the gene;

(b) culturing DFCs (II) under conditions wherein the second Al of the gene is substantially under expd, not expd or over expd and generating a second transcription profile for the cultured cells; and

(c) comparing the first transcription profile with the second transcription profile to identify changes in the levels of gene transcripts;

(10) a NAM (X) comprising a Nt seq encoding a gene product required for proliferation of *Candida albicans* (the gene product comprises a defined amino acid seq given in the specification);

(11) a NAM (XI) comprising a fragment of comprising at least 10 to 100 consecutive Nts; a NAM comprising a Nt seq that hybridizes under stringent condition to a second NAM consisting of:

(a) a Nt seq selected from the seqs given in the specification; or

(b) a Nt seq that encodes a **polypeptide** comprising an amino acid seq defined in the specification (the stringent conditions comprises hybridization to filter-bound DNA in 6 multiply sodium chloride sodium citrate (SSC) at about 45 deg. C followed by one or more washes in 0.2 multiply SSC/0.1 % sodium dodecyl sulfate (SDS) at about 50-65 deg. C);

10/081935

(12) a NAM (XII) obtained from an organism other than *Candida albicans* or *Saccharomyces cerevisiae* comprising a Nt seq having at least 30% identity to a defined seq given in the specification, fragments of at least 25 consecutive Nts, and/or seqs complementary to them as determined using BLAST-N version 2.0 with the default parameters;

(13) a **vector** (XIII) comprising a promoter operably linked to the NAMS above;

(14) a host cell (XIV) containing the vector (XIII);

(15) a polypeptide (XV) comprising a defined amino acid seq given in the specification;

(16) a polypeptide (XVI) obtained from an organism other than *Candida albicans* or *Saccharomyces cerevisiae* comprising an amino acid seq having at least 30% similarity to an amino acid seq defined in the specification as determined using FASTA version 3.0t78 with the default parameters;

(17) a fusion protein (XVII) comprising a fragment of a polypeptide fused to a second polypeptide (the fragment comprises at least 6 consecutive residues of a defined amino acid seq given in the specification;

(18) a method (IXX) of producing a polypeptide, comprising introducing into a cell, a vector comprising a promoter operably linked to a Nt seq encoding a polypeptide comprising a defined amino acid seq given in the specification and culturing the cell so that the Nt seq is expd; and

(19) a method (XX) of producing a polypeptide, comprising providing a cell which comprises a heterologous promoter operably linked to a Nt seq encoding a polypeptide comprising a defined amino acid seq given in the specification and culturing the cell so that the Nt seq is expd.

Note: Other INDEPENDENT CLAIMS are included but have had to be omitted due to lack of space.

USE - The methods are used to identify agents that may be used in the treatment of fungal infections.

Dwg.0/6

L34 ANSWER 9 OF 18 MEDLINE on STN DUPLICATE 3
ACCESSION NUMBER: 2001248085 MEDLINE
DOCUMENT NUMBER: 21189198 PubMed ID: 11292702
TITLE: **Recombinant** urease and urease DNA of **Coccidioides immitis** elicit an immunoprotective response against **coccidioidomycosis** in mice.
AUTHOR: Li K; Yu J J; Hung C Y; Lehmann P F; Cole G T
CORPORATE SOURCE: Department of Microbiology and Immunology, Medical College of Ohio, Toledo, Ohio 43614-5806, USA.
CONTRACT NUMBER: AI19149 (NIAID)
AI37232 (NIAID)
SOURCE: INFECTION AND IMMUNITY, (2001 May) 69 (5) 2878-87.
Journal code: 0246127. ISSN: 0019-9567.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200105
ENTRY DATE: Entered STN: 20010517
Last Updated on STN: 20010517
Entered Medline: 20010510

10/081935

AB **Coccidioides immitis** antigens which stimulate a T helper cell 1 (Th1) pathway of host immune response are considered to be essential components of a vaccine against **coccidioidomycosis**. **Recombinant** urease (rURE) and **recombinant** heat shock **protein** 60 (rHSP60) of *C. immitis* were expressed in *Escherichia coli* and tested as vaccine candidates in BALB/c mice. A synthetic oligodeoxynucleotide which contained unmethylated CpG dinucleotides and was previously shown to enhance a murine Th1 response was used as an immunoadjuvant. T cells isolated from the spleens and lymph nodes of the rURE- and rHSP60-immune mice showed in vitro proliferative responses to the respective **recombinant protein**, but only those T lymphocytes from rURE-immunized mice revealed markedly elevated levels of expression of selected Th1-type cytokine genes. BALB/c mice immunized subcutaneously with rURE and subsequently challenged by the intraperitoneal (i.p.) route with a lethal inoculum of *C. immitis* arthroconidia demonstrated a significant reduction in the level of *C. immitis* infection compared to control animals. rHSP60 was much less effective as a protective antigen. Evaluation of cytokine gene expression in lung tissue and levels of **recombinant** urease-specific immunoglobulins (immunoglobulin G1 [IgG1] versus IgG2a) in murine sera at 12 days after challenge provided additional evidence that immunization with rURE stimulated a Th1 response to the pathogen. Urease was further evaluated by expression of the URE gene in a mammalian plasmid **vector** (pSecTag2A.URE) which was used to immunize mice by the intradermal route. In this case, 82% of the **vector** construct-immunized animals survived more than 40 days after i.p. infection, compared to only 10% of the mice immunized with the **vector** alone. In addition, 87% of the pSecTag2A.URE-immunized survivors had sterile lungs and spleens. These data support the need for further evaluation of the *C. immitis* urease as a candidate vaccine against **coccidioidomycosis**.

L34 ANSWER 10 OF 18 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC. on STN

ACCESSION NUMBER: 2002:223315 BIOSIS
DOCUMENT NUMBER: PREV200200223315
TITLE: Replacement of the gene which encodes a major immunoreactive cell surface antigen and adhesin (SOWgp) of the human fungal pathogen **Coccidioides immitis**.
AUTHOR(S): Hung, C. [Reprint author]; Reichard, U. [Reprint author]; Yu, J. [Reprint author]; Cole, G. T. [Reprint author]
CORPORATE SOURCE: Medical College of Ohio, Toledo, OH, USA
SOURCE: Abstracts of the General Meeting of the American Society for Microbiology, (2001) Vol. 101, pp. 365. print.
Meeting Info.: 101st General Meeting of the American Society for Microbiology. Orlando, FL, USA. May 20-24, 2001. American Society of Microbiology. ISSN: 1060-2011.
DOCUMENT TYPE: Conference; (Meeting)
Conference; Abstract; (Meeting Abstract)
LANGUAGE: English
ENTRY DATE: Entered STN: 3 Apr 2002
Last Updated on STN: 3 Apr 2002

Searcher : Shears 308-4994

10/081935

AB We have previously cloned and expressed a major parasitic cell surface antigen (SOWgp) of *C. immitis* which, upon immunization elicits both antibody and T cell responses in BALB/c mice. Our data also showed that the **recombinant** SOWgp has high affinity for laminin and fibronectin of the extracellular matrix (ECM). To confirm that the SOWgp functions as an adhesin, the SOWgp gene was deleted by a double, locus-specific crossover event. This resulted in replacement of 1.8-kb fragment of the SOWgp gene with a 3.6-kb DNA construct that encodes a hygromycin resistant gene from the pAN7.1 **vector**. Results of PCR analysis, Southern hybridization, and immunoblot assays of parasitic cell homogenates of the Dsowgp mutant compared to the parental strain confirmed that the SOWgp gene was replaced and the antigen was no longer produced. This is the first report of a successful targeted gene knockout of *C. immitis* using the double crossover strategy. Comparative immunoelectron-microscopic examinations of the isolated, crude outer parasitic cell wall fraction of the DELTAsowgp and parental strains showed total loss of patient antibody reactivity with the former. The DELTAsowgp parasitic cells showed 20-50% reduction in their ability to bind to human and murine ECM. Virulence studies of the DELTAsowgp strain in intranasally-challenged BALB/c mice showed 20% survival compared to none in the case of mice challenged with the parental strain. It is apparent that deletion of SOWgp production did not eliminate the ability of the pathogen to cause disease. Preliminary investigations of the influence of **recombinant** SOWgp on immune murine T cells in vitro suggested that the **protein** stimulates a dominant Th2 response, demonstrated by elevated levels of IL-5 production. Our hypothesis is that SOWgp has a major impact on disseminated **coccidioidal** infection by compromising the host immune pathway toward a detrimental Th2 response.

L34 ANSWER 11 OF 18 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC. on STN

ACCESSION NUMBER: 2002:223274 BIOSIS
DOCUMENT NUMBER: PREV200200223274
TITLE: Isolation and characterization of a macrophage binding **protein** of **Coccidioides immitis**.
AUTHOR(S): Cole, G. T. [Reprint author]; Fradin, C. [Reprint author]
CORPORATE SOURCE: Medical College of Ohio, Toledo, OH, USA
SOURCE: Abstracts of the General Meeting of the American Society for Microbiology, (2001) Vol. 101, pp. 356. print.
Meeting Info.: 101st General Meeting of the American Society for Microbiology. Orlando, FL, USA. May 20-24, 2001. American Society of Microbiology. ISSN: 1060-2011.
DOCUMENT TYPE: Conference; (Meeting)
Conference; Abstract; (Meeting Abstract)
LANGUAGE: English
ENTRY DATE: Entered STN: 3 Apr 2002
Last Updated on STN: 3 Apr 2002

AB Early studies have demonstrated the significance of the interplay between **Coccidioides immitis** and macrophages during **coccidioidomycosis**. In this study we report the purification of a ligand expressed at the surface of endospores of

10/081935

C. immitis which binds to an unknown receptor of a murine macrophage line (IC-21). Cell wall **proteins** of different stages of the parasitic cycle were labeled with biotin and subjected to beta-mercaptoethanol extraction. A single macrophage plasma membrane binding **protein** (MBP) with molecular size of 36-kDa was isolated from the endospore extract. The C. immitis MBP was purified and two internal **peptides** of the trypsin-digested fraction were sequenced. Degenerate primers were designed to PCR amplify a fragment of the MBP gene, and the PCR product was then used to screen a C. immitis genomic library for isolation of the full-length gene. The MBP gene is characterized by an open reading frame of 638 bp, and does not contain any introns. The MBP gene was subcloned into the pET28b expression **vector** and the **recombinant**, histidine-tagged fusion **protein** (rMBP) was produced by transformed E. coli strain BL21(DE3). The rMBP was shown to bind to the surface of both IC-21 cells and resident murine peritoneal macrophages. An in vitro study of temporal expression of the MBP gene during the parasitic cycle was conducted by RT-PCR and Northern blot analyses, and showed that maximum expression occurs at the endospore stage. Polyclonal antibody raised in guinea pigs against the rMBP was used for immunoblot analyses of beta-mercaptoethanol extracts of intact parasitic cells, as well as total cell homogenates of the saprobic and parasitic phases of C. immitis. The results showed that the rMBP was present only in beta-mercaptoethanol extracts and homogenates of first generation round cells and endospores. We suggest that the newly-isolated MBP of C. immitis plays a pivotal role in the course of **coccidioid** infections:

L34 ANSWER 12 OF 18 WPIDS COPYRIGHT 2003 THOMSON DERWENT on STN .
ACCESSION NUMBER: 1999-551417 [46] WPIDS
CROSS REFERENCE: 2002-606631 [65]; 2003-829776 [77]
DOC. NO. NON-CPI: N1999-407989
DOC. NO. CPI: C1999-161024
TITLE: Novel chitin-binding fragments of human chitinase
used to treat fungal infections in animals.
DERWENT CLASS: B04 C06 D16 S03
INVENTOR(S): GRAY, P W; TJOELKER, L W
PATENT ASSIGNEE(S): (ICOS-N) ICOS CORP
COUNTRY COUNT: 86
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG

WO 9946390	A1	19990916	(199946)*	EN	83
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC					
MW NL OA PT SD SE SL SZ UG ZW					
W: AE AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES					
FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK					
LR LS LT LU LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG					
SI SK SL TJ TM TR TT UA UG UZ VN YU ZW					
AU 9929989	A	19990927	(200006)		
BR 9908724	A	20001121	(200065)		
NO 2000004522	A	20001101	(200065)		
EP 1078073	A1	20010228	(200113)	EN	
R: AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE					
US 6200951	B1	20010313	(200120)		
CZ 2000003308	A3	20010711	(200147)		

10/081935

HU	2001001078	A2	20010730	(200157)	
SK	2000001354	A3	20010710	(200157)	
MX	2000008863	A1	20010301	(200170)	
JP	2002505882	W	20020226	(200219)	86
CN	1357046	A	20020703	(200265)	
AU	763582	B	20030724	(200355)	

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 9946390	A1	WO 1999-US5343	19990312
AU 9929989	A	AU 1999-29989	19990312
BR 9908724	A	BR 1999-8724	19990312
		WO 1999-US5343	19990312
NO 2000004522	A	WO 1999-US5343	19990312
		NO 2000-4522	20000911
EP 1078073	A1	EP 1999-911320	19990312
		WO 1999-US5343	19990312
US 6200951	B1	US 1998-39198	19980312
CZ 2000003308	A3	WO 1999-US5343	19990312
		CZ 2000-3308	19990312
HU 2001001078	A2	WO 1999-US5343	19990312
		HU 2001-1078	19990312
SK 2000001354	A3	WO 1999-US5343	19990312
		SK 2000-1354	19990312
MX 2000008863	A1	MX 2000-8863	20000911
JP 2002505882	W	WO 1999-US5343	19990312
		JP 2000-535757	19990312
CN 1357046	A	CN 1999-805990	19990312
AU 763582	B	AU 1999-29989	19990312

FILING DETAILS:

PATENT NO	KIND	PATENT NO
AU 9929989	A Based on	WO 9946390
BR 9908724	A Based on	WO 9946390
EP 1078073	A1 Based on	WO 9946390
CZ 2000003308	A3 Based on	WO 9946390
HU 2001001078	A2 Based on	WO 9946390
SK 2000001354	A3 Based on	WO 9946390
JP 2002505882	W Based on	WO 9946390
AU 763582	B Previous Publ.	AU 9929989
	Based on	WO 9946390

PRIORITY APPLN. INFO: US 1998-39198 19980312

AN 1999-551417 [46] WPIDS
CR 2002-606631 [65]; 2003-829776 [77]
AB WO 9946390 A UPAB: 20031128

NOVELTY - Novel human chitinase **polypeptide** and polynucleotide fragments have chitin-binding activity but lack chitinase enzymatic activity.

DETAILED DESCRIPTION - A chitin-binding, chitinase-inactive **polypeptide** (I) comprises a chitin-binding fragment of the 54 C-terminal amino acids (aa) of human chitinase having a 466 aa sequence (S1) (given in the specification).

INDEPENDENT CLAIMS are also included for the following:

10/081935

- (1) a chitin-binding, chitinase-inactive **polypeptide** comprising aa residues X-Y of (S1), where X is a consecutive integer from 347 to 397 and Y is 445;
- (2) a fusion **protein** comprising (I) fused to a heterologous **polypeptide** (e.g. an enzyme);
- (3) a composition comprising (I) and a diluent, optionally further comprising a non-chitinase anti-fungal agent;
- (4) a composition comprising (I) or the fusion **protein** of (1) conjugated to an anti-fungal agent;
- (5) treating fungal infection, comprising administering to a subject suffering from fungal infection the composition of (3) or (4), optionally further comprising administering a non-chitinase anti-fungal agent;
- (6) a composition comprising (I) or the fusion **protein** of (1) conjugated to a detectable label, especially radioisotopes, fluorophores, dyes, electron-dense compounds or enzymes;
- (7) determining the presence of chitin in a sample, comprising contacting the sample with the composition of (6), and determining the amount of labelled **polypeptide** bound to chitin;
- (8) a purified, isolated polynucleotide (especially DNA) encoding (I);
- (9) a **vector** comprising the DNA of (8);
- (10) a host cell stably transformed or transfected with the DNA of (8) in a manner allowing the expression of (I), and
- (11) a monoclonal antibody that specifically binds to an epitope within the 54 C-terminal amino acids of human chitinase having a sequence (S1).

ACTIVITY - Chitin-binding.

MECHANISM OF ACTION - None given.

USE - Chitinase fragments can be used to screen for **proteins** or other molecules that specifically bind to the chitin-binding domain of human chitinase or that modulate its activity. These compounds are useful for immunization, as well as for purifying chitinase, as well as for detection and quantification of chitinase. Polynucleotide fragments of the invention are useful as a source of probes and primers, and to express the **proteins recombinantly**. The chitinase fragments, when conjugated to antifungal compounds, are used to treat animals, especially humans, infected with chitin-containing parasites such as fungi. Fungal infection treated include candidiasis, aspergillosis, **coccidioidomycosis**, blastomycosis, paracoccidioidomycosis, mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis, and dermatophytoses.

ADVANTAGE - Chitin can be degraded by the enzyme chitinase. Use of whole chitinase **protein** for treating infections, especially fungal infections, is problematic. In view of the increasing incidences of life-threatening fungal infection in e.g. immunocompromised individuals, there exists a need for identifying new compounds for treating fungal infection. The chitin-binding fragments of the present invention provide this need.
Dwg.0/0

L34 ANSWER 13 OF 18

MEDLINE on STN

DUPLICATE 4

ACCESSION NUMBER: 1999085554 MEDLINE

DOCUMENT NUMBER: 99085554 PubMed ID: 9868663

TITLE: Detecting serum antibodies to a purified
recombinant proline-rich antigen of
Coccidioides immitis in patients with

10/081935

coccidioidomycosis.
AUTHOR: Orsborn K I; Galgiani J N
CORPORATE SOURCE: Research and Medical Services, Veterans Affairs
Medical Center, Tucson, Arizona, USA.
SOURCE: CLINICAL INFECTIOUS DISEASES, (1998 Dec) 27 (6)
1475-8.
Journal code: 9203213. ISSN: 1058-4838.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199903
ENTRY DATE: Entered STN: 19990324
Last Updated on STN: 19990324
Entered Medline: 19990308

AB In previous work, antibodies in serum samples from patients with **coccidioidomycosis** were found to react with a proline-rich antigen (PRA) isolated from spherules of **Coccidioides immitis**, and the gene encoding this antigen was cloned. We expressed and purified **recombinant** PRA (rPRA) by removing the majority of amino acids contributed by the **vector** from the fusion **protein**. Purified rPRA reacted with serum IgG antibodies in 37 of 42 patients with culture-proven progressive pulmonary or extrapulmonary **coccidioidal** disease; specific antibodies in dilutions ranging from 1:40 to 1:102,400 were demonstrated (sensitivity, 88%). In contrast, for > 95% of patients without **coccidioidomycosis** reactivity of < 1:40 was demonstrated (specificity, 97%). Of 18 patients with primary self-limited **coccidioidomycosis**, none had detectable antibodies in serum samples collected up to 141 days after illness began. The association of antibodies to rPRA with progressive infection may have prognostic value.

L34 ANSWER 14 OF 18 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC. on STN

ACCESSION NUMBER: 1997:500830 BIOSIS
DOCUMENT NUMBER: PREV199799800033
TITLE: Isolation and characterization of the urease gene (URE) from the pathogenic fungus **Coccidioides immitis**.
AUTHOR(S): Yu, Jieh-Juen; Smithson, S. L.; Thomas, Pei W.; Kirkland, Theo N.; Cole, Garry T. [Reprint author]
CORPORATE SOURCE: Dep. Microbiol. Immunol., Med. Coll. Ohio, 3000 Arlington Ave., Toledo, OH 43699, USA
SOURCE: Gene (Amsterdam), (1997) Vol. 198, No. 1-2, pp. 387-391.
CODEN: GENED6. ISSN: 0378-1119.
DOCUMENT TYPE: Article
LANGUAGE: English
ENTRY DATE: Entered STN: 21 Nov 1997
Last Updated on STN: 21 Nov 1997

AB The urease (URE)-encoding gene from **Coccidioides immitis** (Ci), a respiratory fungal pathogen of humans, was cloned, sequenced, chromosome-mapped and expressed. Both the genomic and cDNA sequences are reported. The transcription start point and poly(A)-addition site were confirmed. The URE gene contains eight introns and a 2517-bp ORF that translates a 839-amino-acid (aa) **protein** of 91.5 kDa and pI of 5.5, as deduced by computer

10/081935

analysis of the nucleotide sequence. The translated **protein** revealed eight putative N-glycosylation sites. The deduced URE showed comparable levels of homology to reported URE of the jack bean plant (*Canavalia ensiformis*; 71.8%) and URE of several genera of bacteria (Bp, 71.7%; Hp, 68.3%; Ka, 71.6%; Pm, 71.9%). The URE gene was mapped to chromosome III of Ci and was shown to be a single copy gene by Southern hybridization. Expression of a 1687-bp fragment of the URE gene in *E. coli* resulted in the production of a 63-kDa **recombinant protein** that was recognized in an immunoblot by antiserum raised against the Ka URE homolog. This is the first report of a fungal URE gene.

L34 ANSWER 15 OF 18 MEDLINE on STN DUPLICATE 5
ACCESSION NUMBER: 96294782 MEDLINE
DOCUMENT NUMBER: 96294782 PubMed ID: 8698497
TITLE: Molecular cloning and characterization of **Coccidioides immitis** antigen 2 cDNA.
AUTHOR: Zhu Y; Yang C; Magee D M; Cox R A
CORPORATE SOURCE: Department of Clinical Investigation, Texas Center for Infectious Disease, San Antonio 78223, USA.
CONTRACT NUMBER: AI2143 (NIAID)
SOURCE: INFECTION AND IMMUNITY, (1996 Jul) 64 (7) 2695-9.
Journal code: 0246127. ISSN: 0019-9567.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
OTHER SOURCE: GENBANK-U32518
ENTRY MONTH: 199609
ENTRY DATE: Entered STN: 19960912
Last Updated on STN: 19960912
Entered Medline: 19960904

AB Previous experiments have provided evidence that **Coccidioides immitis** antigen 2 (Ag2) is a major T-cell-reactive component of mycelia and spherule cell walls. Here we report the identification and cloning of the cDNA that encodes Ag2 from a lambda ZAP cDNA expression library constructed from spherule-derived RNA. DNA sequence analysis established that the 1,255-bp clone contains a 174-bp 5' untranslated region, a 582-bp open reading frame which encodes for a **protein** consisting of 194 amino acids, and a 375-bp 3' untranslated region, including a poly(A) tail. The **recombinant Ag2 protein** has a predicted molecular mass of 19.5 kDa and contains an 18-amino-acid N terminus which has been tentatively identified as a signal **peptide**. The Ag2 cDNA was ligated into the pGEX-4T-3 **vector** and expressed in *Escherichia coli* TG-1 cells as a glutathione S-transferase fusion **protein**. The **recombinant** fusion **protein** showed reactivity with sera from patients with **coccidioidomycosis** and elicited delayed-type footpad hypersensitivity responses in **Coccidioides**-immune mice. These results suggest that the Ag2 cDNA can be used for the large-scale production of this immunologically important **protein**.

L34 ANSWER 16 OF 18 MEDLINE on STN DUPLICATE 6
ACCESSION NUMBER: 96239008 MEDLINE
DOCUMENT NUMBER: 96239008 PubMed ID: 8675298
TITLE: Molecular cloning and characterization of the

10/081935

Coccidioides immitis complement
fixation/chitinase antigen.
AUTHOR: Yang C; Zhu Y; Magee D M; Cox R A
CORPORATE SOURCE: Department of Clinical Investigation, Texas Center
for Infectious Diseases, San Antonio, 78223, USA.
CONTRACT NUMBER: AI 12431 (NIAID)
SOURCE: INFECTION AND IMMUNITY, (1996 Jun) 64 (6) 1992-7.
Journal code: 0246127. ISSN: 0019-9567.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
OTHER SOURCE: GENBANK-U33265
ENTRY MONTH: 199608
ENTRY DATE: Entered STN: 19960822
Last Updated on STN: 19960919
Entered Medline: 19960812

AB Detection of anti-**Coccidioides** complement-fixing (CF)
antibody is a valuable diagnostic and prognostic aid in
coccidioidomycosis. The CF antibody response is directed
against a heat-labile antigen that has chitinase activity, hereafter
referred to as the CF/chitinase **protein**. To identify and
clone this immunoreactive enzyme, we constructed a
Coccidioides immitis cDNA lambda ZAP expression library from
spherule RNA and detected fusion **peptides** expressing CF
epitopes by immunoscreening. A cDNA clone consisting of 1,623 bp
was identified, sequenced, and found to contain a single open
reading frame that encodes a **protein** of 47 kDa with 427
amino acids. Deduced amino acid sequence analyses showed that the
cloned CF/chitinase cDNA contains a 35-amino-acid region, beginning
at Ser-18 and ending at and ending at Arg-52 which has 92% homology
with the reported N-terminal amino acid sequence of authentic
CF/chitinase **protein**. The first 17 amino acids in the
deduced sequence of the cloned cDNA are not present on the mature
CF/chitinase **protein**, suggesting that it may be a signal
peptide. Expression of the CF/chitinase cDNA insert by
using the pGEX-4T-3 **vector** yields a fusion **peptide**
that bears CF-specific epitopes and shows chitinase activity. The
CF/chitinase clone will enable large-scale production of the
recombinant CF antigen for use in immunoassays and
facilitate studies on the role of chitinase in the morphogenesis of
C. immitis.

L34 ANSWER 17 OF 18 MEDLINE on STN DUPLICATE 7
ACCESSION NUMBER: 96009757 MEDLINE
DOCUMENT NUMBER: 96009757 PubMed ID: 7558310
TITLE: Molecular and biochemical characterization of a
Coccidioides immitis-specific antigen.
AUTHOR: Pan S; Cole G T
CORPORATE SOURCE: Department of Microbiology, Medical College of Ohio,
Toledo 43699-0008, USA.
CONTRACT NUMBER: AI19149 (NIAID)
SOURCE: INFECTION AND IMMUNITY, (1995 Oct) 63 (10) 3994-4002.
Journal code: 0246127. ISSN: 0019-9567.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals

Searcher : Shears 308-4994

10/081935

OTHER SOURCE: GENBANK-L36551
ENTRY MONTH: 199510
ENTRY DATE: Entered STN: 19951227
Last Updated on STN: 19951227
Entered Medline: 19951030

AB Results of earlier investigations have indicated that the saprobic phase of **Coccidioides immitis** produces a heat-stable, 19-kDa antigen with serine proteinase activity which has been suggested to be specific for this pathogenic fungus. In the present study we have determined the N-terminal and partial internal amino acid sequences of the purified, 19-kDa antigen, cloned the gene which encodes this **polypeptide**, and confirmed that the secreted proteinase is a **Coccidioides**-specific antigen (CS-Ag). Both the genomic and cDNA sequences are reported and reveal that the *csa* gene which encodes this antigen has no introns. A 543-bp open reading frame encodes a 181-amino-acid-containing **protein** with a predicted molecular mass of 19.8 kDa and an isoelectric point of 8.3. The *csa* gene was localized on chromosome I of three representative *C. immitis* clinical isolates on the basis of Southern hybridizations. Expression of the *csa* gene in *Escherichia coli* using the pET21a plasmid **vector** yielded a **recombinant protein** that was recognized in immunoblot assays by antibody raised to the purified 19-kDa CS-Ag. Secretion of the native antigen is suggested to occur by cleavage of a putative 23-residue signal **peptide**. The native CS-Ag showed a low degree of glycosylation. Analysis of the carbohydrate composition of the CS-Ag revealed xylose, mannose, galactose, and glucose. However, the purified antigen showed no affinity for concanavalin A. A PCR method with specificity and high sensitivity for detection of *C. immitis* genomic DNA, using a pair of synthetic oligonucleotide primers whose sequences were based on that of the *csa* gene, was developed. A 520-bp product was amplified only when *C. immitis* genomic DNA was used as the template. The lower limits of DNA detection using this PCR method were 1 pg of *C. immitis* genomic DNA by ethidium bromide staining and 100 fg after Southern hybridization. The *csa* gene-based PCR method for detection of *C. immitis* DNA is useful for culture identification and may have clinical applications for the diagnosis of **coccidioidal** infections.

L34 ANSWER 18 OF 18 MEDLINE on STN
ACCESSION NUMBER: 94129182 MEDLINE
DOCUMENT NUMBER: 94129182 PubMed ID: 8298279
TITLE: Identification of antigens of **Coccidioides immitis** which stimulated immune T lymphocytes.
AUTHOR: Cole G T; Kirkland T N
CORPORATE SOURCE: Department of Botany, University of Texas, Austin 78713.
CONTRACT NUMBER: AI19149 (NIAID)
SOURCE: ARCHIVES OF MEDICAL RESEARCH, (1993 Autumn) 24 (3) 281-91.
Journal code: 9312706. ISSN: 0188-4409.
PUB. COUNTRY: Mexico
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199403
ENTRY DATE: Entered STN: 19940318

10/081935

Last Updated on STN: 19940318

Entered Medline: 19940308

AB T-cell mediated immune response to **coccidioidomycosis** has been shown to be the principal mechanism of resistance to this respiratory fungal disease in experimental animals. In this study, a **Coccidioides immitis** antigen-specific murine T-cell line was used to identify macromolecules capable of eliciting an immune mouse T-cell proliferative response. The murine T-cell line was selected on the basis of its strong positive response to a soluble conidial wall fraction (SCWF), which had previously been shown to be reactive in humoral and cellular immunoassays. An antigen-specific T-cell line rather than T-cell clones was used to identify multiple antigens. The T-cell immunoblot method was employed first to identify immunoreactive sub-fractions of the SCWF, and then to identify T-cell fusion **proteins** (FPs) obtained from a cDNA expression library constructed in lambda gt11. The library was screened with anti-SCWF. The nucleotide sequence of a 0.2 kilobase cDNA insert encoding a FP which elicited vigorous T-cell response was determined. A construct of this insert was subcloned into the pET expression **vector** system and 6.5-kilodalton (kDa) **recombinant protein** (RP) expressed in *Escherichia coli* was isolated. The RP and FP were shown to be homologous on the basis of identify of their amino acid sequences. Antibody raised in guinea pigs against the RP recognized a 59-kDa native **protein** of the mycelial culture filtrate produced by three separate strains of *C. immitis*, and reacted with the cell wall of arthroconidia as detected by immunofluorescence microscopy. In this study we have identified and partly characterized a potentially important T-cell stimulating antigen of *C. immitis*.

(FILE 'MEDLINE' ENTERED AT 09:35:53 ON 04 DEC 2003)

L35 720 SEA FILE=MEDLINE ABB=ON PLU=ON COCCIDIOIDES/CT
L36 120410 SEA FILE=MEDLINE ABB=ON PLU=ON PROTEINS/CT
L37 86 SEA FILE=MEDLINE ABB=ON PLU=ON POLYPROTEINS/CT
L38 73063 SEA FILE=MEDLINE ABB=ON PLU=ON PEPTIDES/CT
L39 7 SEA FILE=MEDLINE ABB=ON PLU=ON L35 AND (L36 OR L37 OR L38)

L39 ANSWER 1 OF 7 MEDLINE on STN

AN 2002322678 MEDLINE

TI Localization within a proline-rich antigen (Ag2/PRA) of protective antigenicity against infection with *Coccidioides immitis* in mice.

AU Peng Tao; Shubitz Lisa; Simons Julie; Perrill Robert; Orsborn Kris I; Galgiani John N

SO INFECTION AND IMMUNITY, (2002 Jul) 70 (7) 3330-5.

Journal code: 0246127. ISSN: 0019-9567.

AB Subunits of a proline-rich coccidioidal antigen (Ag2/PRA) of *Coccidioides immitis* were analyzed by comparison as vaccines in mice. The optimal dose of plasmid vaccine encoding full-length Ag2/PRA was determined to be between 10 and 100 microg. Mice vaccinated with plasmids encoding amino acids (aa) 1 to 106 were as protective as full-length Ag2/PRA (aa 1 to 194). The subunit from aa 27 to 106 was significantly but less protective. Plasmids encoding aa 90 to 151 or aa 90 to 194 were not protective. Analogous results were obtained with recombinant vaccines of the same amino acid sequences. In addition, mixtures of aa 90 to 194 with either aa 1 to 106 or aa 27 to 106 did not enhance protection compared to the active single-recombinant subunits alone. Humoral

10/081935

response of total immunoglobulin G (IgG) and subclasses IgG1 and IgG2a were detectable in subunit vaccinations but at significantly (100-fold) lower concentrations than after vaccination with plasmids encoding full-length Ag2/PRA. Since virtually all protection by vaccination with full-length Ag2/PRA can be accounted for in the first half of the protein (aa 1 to 106), this subunit could make a multicomponent vaccine more feasible by reducing the quantity of protein per dose and the possibility of an untoward reactions to a foreign protein.

- L39 ANSWER 2 OF 7 MEDLINE on STN
AN 1999094977 MEDLINE
TI Proline-rich vaccine candidate antigen of *Coccidioides immitis*: conservation among isolates and differential expression with spherule maturation.
AU Peng T; Orsborn K I; Orbach M J; Galgiani J N
SO JOURNAL OF INFECTIOUS DISEASES, (1999 Feb) 179 (2) 518-21.
Journal code: 0413675. ISSN: 0022-1899.
AB A proline rich antigen (PRA), which protects mice against *Coccidioides immitis*, has been analyzed for differential expression and variation among isolates. Northern blots of RNA from different stages of growth were probed with previously cloned cDNA and showed that mRNA for PRA increased as spherules transformed and matured from mycelia. This pattern corresponds to the relative potency of whole cell vaccines from similar preparations. The PRA gene was then cloned from a genomic library of the Silveira strain of *C. immitis* and its sequence analyzed. Eight other coccidioidal isolates, selected for diversity in geographic origin and resulting clinical disease, demonstrated heterogeneity in Southern blots and in sequences of polymerase chain reaction products. Silveira differed from other California isolates at 33 of 555 bases, whereas it differed from non-California isolates by ≤ 2 bases. Only one of these substitutions affected protein sequence. Thus, tests or vaccines based on this gene are likely to cover most isolates.
- L39 ANSWER 3 OF 7 MEDLINE on STN
AN 96144702 MEDLINE
TI Cloning and sequence analysis of the cDNA for a protein from *Coccidioides immitis* with immunogenic potential.
AU Dugger K O; Villareal K M; Ngyuen A; Zimmermann C R; Law J H; Galgiani J N
SO BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, (1996 Jan 17) 218 (2) 485-9.
Journal code: 0372516. ISSN: 0006-291X.
AB We have cloned and sequenced the cDNA encoding an immunoreactive protein from the pathogenic fungus *Coccidioides immitis* which stimulates human T cells and has been associated with protective vaccines in mice. The transcript contained an open reading frame encoding 194 amino acids with a calculated molecular weight of 19.5 kDa, a 151 base 5' untranslated region (UTR), and a 468 base 3'UTR. A four member repeat motif, usually thr-ala-glu-pro, exists for amino acids 98 through 141. Deduced amino acid sequence derived from the cDNA was identical with previously determined internal amino acid sequence from the native protein, and goat antiserum raised against the purified fungal protein reacted with an inducible fusion protein translated from this cDNA. Using this cDNA to produce recombinant protein will allow direct testing of its role in human immunity to coccidioidomycosis and may lead to new diagnostic

10/081935

tests.

L39 ANSWER 4 OF 7 MEDLINE on STN
AN 90383461 MEDLINE
TI Activity of cilofungin against *Coccidioides immitis*: differential in vitro effects on mycelia and spherules correlated with in vivo studies.
AU Galgiani J N; Sun S H; Clemons K V; Stevens D A
SO JOURNAL OF INFECTIOUS DISEASES, (1990 Oct) 162 (4) 944-8.
Journal code: 0413675. ISSN: 0022-1899.
AB Cilofungin, a new antifungal drug, was found to inhibit mycelial growth of *Coccidioides immitis*. Light and electron microscopic observations indicated delay in development of the outer hyphal wall. Cilofungin also blocked incorporation of the chitin substrate, N-acetylglucosamine, into mycelia. However, when *C. immitis* was grown under conditions that induced spherule development, drug effects were dramatically decreased. Furthermore, efficacy of cilofungin in treatment of murine coccidioidomycosis could not be demonstrated. These studies indicate that glucan-synthase inhibitors have activity against *C. immitis*, and other compounds with different pharmacologic properties or in combination with other antifungal drugs may exploit this biologic effect.

L39 ANSWER 5 OF 7 MEDLINE on STN
AN 75150364 MEDLINE
TI Immunocompetence and prognosis in patients with gynecologic cancer.
AU Nalick R H; Disaia P J; Rea T H; Morrow C P
SO GYNECOLOGIC ONCOLOGY, (1974 Apr) 2 (1) 81-92.
Journal code: 0365304. ISSN: 0090-8258.

L39 ANSWER 6 OF 7 MEDLINE on STN
AN 72041293 MEDLINE
TI Fractionation and composition studies of skin test-active components of sensitins from *Coccidioides immitis*.
AU Anderson K L; Wheat R W; Conant N F
SO APPLIED MICROBIOLOGY, (1971 Sep) 22 (3) 294-9.
Journal code: 7605802. ISSN: 0003-6919.

L39 ANSWER 7 OF 7 MEDLINE on STN
AN 68239308 MEDLINE
TI Sarcoidosis. II. A clinical evaluation of the alteration in delayed hypersensitivity.
AU Lordon R E; Young R L; Shapiro S S; Smith R E; Weg J G
SO AMERICAN REVIEW OF RESPIRATORY DISEASE, (1968 Jun) 97 (6) 1009-16.
Journal code: 0370523. ISSN: 0003-0805.

(FILE 'HCAPLUS' ENTERED AT 09:41:53 ON 04 DEC 2003)

L52 208 SEA FILE=HCAPLUS ABB=ON PLU=ON C IMMITIS
L53 74 SEA FILE=HCAPLUS ABB=ON PLU=ON L52 AND (PROTEIN OR POLYPEPTIDE OR PROTEIN OR PEPTIDE)
L54 10 SEA FILE=HCAPLUS ABB=ON PLU=ON L53 AND VECTOR
L55 7 SEA FILE=HCAPLUS ABB=ON PLU=ON L54 AND RECOMBIN?

L56 0 L55 NOT L32

(FILE 'MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH,

10/081935

JICST-EPLUS, JAPIO' ENTERED AT 09:43:13 ON 04 DEC 2003)

L57 22 S L55
L58 0 S L57 NOT L33

(FILE 'USPATFULL' ENTERED AT 09:44:23 ON 04 DEC 2003)

L59 635 SEA FILE=USPATFULL ABB=ON PLU=ON (COCCIDIOD? OR
C) (W) IMMITIS OR COCCIDIOIDOMYC?
L60 4 SEA FILE=USPATFULL ABB=ON PLU=ON L59(10A) (PROTEIN OR
PEPTIDE OR POLYPEPTIDE OR POLYPROTEIN)

L59 635 SEA FILE=USPATFULL ABB=ON PLU=ON (COCCIDIOD? OR
C) (W) IMMITIS OR COCCIDIOIDOMYC?
L65 246 SEA FILE=USPATFULL ABB=ON PLU=ON L59(S) (PROTEIN OR
PEPTIDE OR POLYPEPTIDE OR POLYPROTEIN)
L66 2 SEA FILE=USPATFULL ABB=ON PLU=ON L65(S) (VECTOR(3A) RECOM
BIN?)

L67 6 L60 OR L66

L67 ANSWER 1 OF 6 USPATFULL on STN

ACCESSION NUMBER: 2003:311860 USPATFULL
TITLE: Glucanoyl transferase-1 protein useful for
immunization against Coccidioides spp.
INVENTOR(S): Cole, Garry T., Toledo, OH, UNITED STATES
Delgado, Nelson, Delran, NJ, UNITED STATES
Yu, Jieh-Juen, Toledo, OH, UNITED STATES
Xue, Jianmin, Toledo, OH, UNITED STATES

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2003219455	A1	20031127
APPLICATION INFO.:	US 2003-417997	A1	20030416 (10)

	NUMBER	DATE
PRIORITY INFORMATION:	US 2002-374152P	20020422 (60)
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	Michael R. Ward, Morrison & Foerster LLP, 425 Market Street, San Francisco, CA, 94105-2482	
NUMBER OF CLAIMS:	52	
EXEMPLARY CLAIM:	1	
NUMBER OF DRAWINGS:	4 Drawing Page(s)	
LINE COUNT:	2165	

AB The present invention provides compositions of
1,3- β -glucanoyltransferase (Gel-1) antigens and
polynucleotides encoding the Gel-1 antigens, which Gel-1 antigens
are useful for generating an immunological response in an
individual and in therapeutic and diagnostic applications of
infections due to pathogenic Coccidioides spp. fungi, such as C.
immitis or C. posadasii.

INCL INCLM: 424/190.100
INCLS: 435/069.300; 435/070.210; 435/252.300; 435/320.100;
530/350.000; 536/023.700

Searcher : Shears 308-4994

10/081935

NCL NCLM: 424/190.100
NCLS: 435/069.300; 435/070.210; 435/252.300; 435/320.100;
530/350.000; 536/023.700

L67 ANSWER 2 OF 6 USPATFULL on STN

ACCESSION NUMBER: 2003:258639 USPATFULL

TITLE: 207 human secreted proteins

INVENTOR(S): Ni, Jian, Germantown, MD, UNITED STATES
Ebner, Reinhard, Gaithersburg, MD, UNITED STATES
LaFleur, David W., Washington, DC, UNITED STATES
Moore, Paul A., Germantown, MD, UNITED STATES
Olsen, Henrik S., Gaithersburg, MD, UNITED STATES
Rosen, Craig A., Laytonsville, MD, UNITED STATES
Ruben, Steven M., Olney, MD, UNITED STATES
Soppet, Daniel R., Centreville, VA, UNITED STATES
Young, Paul E., Gaithersburg, MD, UNITED STATES
Shi, Yanggu, Gaithersburg, MD, UNITED STATES
Florence, Kimberly A., Rockville, MD, UNITED STATES
Wei, Ying-Fei, Berkeley, CA, UNITED STATES
Florence, Charles, Rockville, MD, UNITED STATES
Hu, Jing-Shan, Mountain View, CA, UNITED STATES
Li, Yi, Sunnyvale, CA, UNITED STATES
Kyaw, Hla, Frederick, MD, UNITED STATES
Fischer, Carrie L., Burke, VA, UNITED STATES
Ferrie, Ann M., Painted Post, NY, UNITED STATES
Fan, Ping, Potomac, MD, UNITED STATES
Feng, Ping, Gaithersburg, MD, UNITED STATES
Endress, Gregory A., Florence, MA, UNITED STATES
Dillon, Patrick J., Carlsbad, CA, UNITED STATES
Carter, Kenneth C., North Potomac, MD, UNITED STATES
Brewer, Laurie A., St. Paul, MN, UNITED STATES
Yu, Guo-Liang, Berkeley, CA, UNITED STATES
Zeng, Zhizhen, Lansdale, PA, UNITED STATES
Greene, John M., Gaithersburg, MD, UNITED STATES

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2003181692	A1	20030925
APPLICATION INFO.:	US 2001-933767	A1	20010822 (9)
RELATED APPLN. INFO.:	Continuation-in-part of Ser. No. WO 2001-US5614, filed on 21 Feb 2001, PENDING Continuation-in-part of Ser. No. US 1998-205258, filed on 4 Dec 1998, PENDING		

	NUMBER	DATE
PRIORITY INFORMATION:	US 2000-184836P	20000224 (60)
	US 2000-193170P	20000329 (60)
	US 1997-48885P	19970606 (60)
	US 1997-49375P	19970606 (60)
	US 1997-48881P	19970606 (60)
	US 1997-48880P	19970606 (60)
	US 1997-48896P	19970606 (60)
	US 1997-49020P	19970606 (60)
	US 1997-48876P	19970606 (60)
	US 1997-48895P	19970606 (60)

Searcher : Shears 308-4994

10/081935

US 1997-48884P	19970606 (60)
US 1997-48894P	19970606 (60)
US 1997-48971P	19970606 (60)
US 1997-48964P	19970606 (60)
US 1997-48882P	19970606 (60)
US 1997-48899P	19970606 (60)
US 1997-48893P	19970606 (60)
US 1997-48900P	19970606 (60)
US 1997-48901P	19970606 (60)
US 1997-48892P	19970606 (60)
US 1997-48915P	19970606 (60)
US 1997-49019P	19970606 (60)
US 1997-48970P	19970606 (60)
US 1997-48972P	19970606 (60)
US 1997-48916P	19970606 (60)
US 1997-49373P	19970606 (60)
US 1997-48875P	19970606 (60)
US 1997-49374P	19970606 (60)
US 1997-48917P	19970606 (60)
US 1997-48949P	19970606 (60)
US 1997-48974P	19970606 (60)
US 1997-48883P	19970606 (60)
US 1997-48897P	19970606 (60)
US 1997-48898P	19970606 (60)
US 1997-48962P	19970606 (60)
US 1997-48963P	19970606 (60)
US 1997-48877P	19970606 (60)
US 1997-48878P	19970606 (60)
US 1997-57645P	19970905 (60)
US 1997-57642P	19970905 (60)
US 1997-57668P	19970905 (60)
US 1997-57635P	19970905 (60)
US 1997-57627P	19970905 (60)
US 1997-57667P	19970905 (60)
US 1997-57666P	19970905 (60)
US 1997-57764P	19970905 (60)
US 1997-57643P	19970905 (60)
US 1997-57769P	19970905 (60)
US 1997-57763P	19970905 (60)
US 1997-57650P	19970905 (60)
US 1997-57584P	19970905 (60)
US 1997-57647P	19970905 (60)
US 1997-57661P	19970905 (60)
US 1997-57662P	19970905 (60)
US 1997-57646P	19970905 (60)
US 1997-57654P	19970905 (60)
US 1997-57651P	19970905 (60)
US 1997-57644P	19970905 (60)
US 1997-57765P	19970905 (60)
US 1997-57762P	19970905 (60)
US 1997-57775P	19970905 (60)
US 1997-57648P	19970905 (60)
US 1997-57774P	19970905 (60)
US 1997-57649P	19970905 (60)
US 1997-57770P	19970905 (60)
US 1997-57771P	19970905 (60)
US 1997-57761P	19970905 (60)
US 1997-57760P	19970905 (60)

10/081935

US 1997-57776P 19970905 (60)
US 1997-57778P 19970905 (60)
US 1997-57629P 19970905 (60)
US 1997-57628P 19970905 (60)
US 1997-57777P 19970905 (60)
US 1997-57634P 19970905 (60)
US 1997-70923P 19971218 (60)
US 1998-92921P 19980715 (60)
US 1998-94657P 19980730 (60)
US 1997-70923P 19971218 (60)
US 1998-92921P 19980715 (60)
US 1998-94657P 19980730 (60)

DOCUMENT TYPE: Utility
FILE SEGMENT: APPLICATION
LEGAL REPRESENTATIVE: HUMAN GENOME SCIENCES INC, 9410 KEY WEST AVENUE,
ROCKVILLE, MD, 20850
NUMBER OF CLAIMS: 23
EXEMPLARY CLAIM: 1
NUMBER OF DRAWINGS: 10 Drawing Page(s)
LINE COUNT: 32746

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating diseases, disorders, and/or conditions related to these novel human secreted proteins.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

INCL INCLM: 536/023.100
INCLS: 530/350.000; 435/325.000; 435/183.000; 435/069.100;
435/320.100
NCL NCLM: 536/023.100
NCLS: 530/350.000; 435/325.000; 435/183.000; 435/069.100;
435/320.100

L67 ANSWER 3 OF 6 USPATFULL on STN

ACCESSION NUMBER: 2003:127198 USPATFULL
TITLE: Death associated kinase containing ankyr in repeats (DAKAR) and methods of use
INVENTOR(S): Bird, Timothy A., Bainbridge Island, WA, UNITED STATES
Holland, Pamela M., Seattle, WA, UNITED STATES
Peschon, Jacques J., Seattle, WA, UNITED STATES
Virca, George D., Bellevue, WA, UNITED STATES

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2003087411	A1	20030508
APPLICATION INFO.:	US 2002-164080	A1	20020604 (10)

	NUMBER	DATE
PRIORITY INFORMATION:	US 2001-295959P	20010604 (60)
	US 2001-334362P	20011129 (60)

DOCUMENT TYPE: Utility

10/081935

FILE SEGMENT: APPLICATION
LEGAL REPRESENTATIVE: IMMUNEX CORPORATION, LAW DEPARTMENT, 51
UNIVERSITY STREET, SEATTLE, WA, 98101
NUMBER OF CLAIMS: 16
EXEMPLARY CLAIM: 1
NUMBER OF DRAWINGS: 4 Drawing Page(s)
LINE COUNT: 5574

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB This invention relates to DAKAR, a new member of the serine/threonine kinase family, methods of making such polypeptides, and to methods of using them to treat conditions associated with apoptosis and epithelial proliferation and differentiation, as well as methods to identify compounds that alter DAKAR-associated cellular activities.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

INCL INCLM: 435/194.000
INCLS: 435/069.100; 435/320.100; 435/325.000; 536/023.200
NCL NCLM: 435/194.000
NCLS: 435/069.100; 435/320.100; 435/325.000; 536/023.200

L67 ANSWER 4 OF 6 USPATFULL on STN

ACCESSION NUMBER: 2003:38356 USPATFULL
TITLE: 125 human secreted proteins
INVENTOR(S): Rosen, Craig A., Laytonsville, MD, UNITED STATES
Feng, Ping, Gaithersburg, MD, UNITED STATES
Ruben, Steven M., Olney, MD, UNITED STATES
Ebner, Reinhard, Gaithersburg, MD, UNITED STATES
Olsen, Henrik S., Gaithersburg, MD, UNITED STATES
Ni, Jian, Germantown, MD, UNITED STATES
Wei, Ying-Fei, Berkeley, CA, UNITED STATES
Soppet, Daniel R., Centreville, VA, UNITED STATES
Moore, Paul A., Germantown, MD, UNITED STATES
Kyaw, Hla, Frederick, MD, UNITED STATES
LaFleur, David W., Washington, DC, UNITED STATES
Shi, Yanggu, Gaithersburg, MD, UNITED STATES
Janat, Fouad, Westerly, RI, UNITED STATES
Endress, Gregory A., Florence, MA, UNITED STATES
Carter, Kenneth C., North Potomac, MD, UNITED STATES
Birse, Charles E., North Potomac, MD, UNITED STATES

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2003028003	A1	20030206
APPLICATION INFO.:	US 2001-974879	A1	20011012 (9)
RELATED APPLN. INFO.:	Continuation-in-part of Ser. No. US 2001-818683, filed on 28 Mar 2001, PENDING Continuation of Ser. No. US 1999-305736, filed on 5 May 1999, PENDING Continuation-in-part of Ser. No. WO 1998-US23435, filed on 4 Nov 1998, UNKNOWN		

	NUMBER	DATE
PRIORITY INFORMATION:	US 2000-239893P	20001013 (60)
	US 1997-64911P	19971107 (60)
	US 1997-64912P	19971107 (60)

Searcher : Shears 308-4994

10/081935

US 1997-64983P 19971107 (60)
US 1997-64900P 19971107 (60)
US 1997-64988P 19971107 (60)
US 1997-64987P 19971107 (60)
US 1997-64908P 19971107 (60)
US 1997-64984P 19971107 (60)
US 1997-64985P 19971107 (60)
US 1997-66094P 19971117 (60)
US 1997-66100P 19971117 (60)
US 1997-66089P 19971117 (60)
US 1997-66095P 19971117 (60)
US 1997-66090P 19971117 (60)

DOCUMENT TYPE: Utility
FILE SEGMENT: APPLICATION
LEGAL REPRESENTATIVE: HUMAN GENOME SCIENCES INC, 9410 KEY WEST AVENUE,
ROCKVILLE, MD, 20850

NUMBER OF CLAIMS: 23
EXEMPLARY CLAIM: 1
NUMBER OF DRAWINGS: 3 Drawing Page(s)
LINE COUNT: 36277

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating diseases, disorders, and/or conditions related to these novel human secreted proteins.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

INCL INCLM: 536/023.200
INCLS: 435/069.100; 435/325.000; 435/320.100; 435/183.000
NCL NCLM: 536/023.200
NCLS: 435/069.100; 435/325.000; 435/320.100; 435/183.000

L67 ANSWER 5 OF 6 USPATFULL on STN

ACCESSION NUMBER: 2002:198286 USPATFULL
TITLE: rSOWgp58 protein and its use as an antigen
INVENTOR(S): Hung, Chiung-Yu, Toledo, OH, UNITED STATES
Cole, Garry T., Toledo, OH, UNITED STATES
PATENT ASSIGNEE(S): Medical College of Ohio (U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2002106380	A1	20020808
APPLICATION INFO.:	US 2001-850677	A1	20010507 (9)

	NUMBER	DATE
PRIORITY INFORMATION:	US 2000-202754P	20000508 (60)
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	EMCH, SCHAFFER, SCHAUB & PORCELLO CO, P O BOX 916, ONE SEAGATE SUITE 1980, TOLEDO, OH, 43697	
NUMBER OF CLAIMS:	32	
EXEMPLARY CLAIM:	1	
NUMBER OF DRAWINGS:	6 Drawing Page(s)	

Searcher : Shears 308-4994

10/081935

LINE COUNT: 1476

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB The present invention provides bacterial plasmids and recombinant SOW proteins that are useful as antigens for serodiagnosis of coccidiomycosis.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

INCL INCLM: 424/191.100

INCLS: 435/007.220; 530/395.000

NCL NCLM: 424/191.100

NCLS: 435/007.220; 530/395.000

L67 ANSWER 6 OF 6 USPATFULL on STN

ACCESSION NUMBER: 2001:208993 USPATFULL

TITLE: Candida albicans gene (CSA1) encoding a mycelial surface antigen, and uses thereof

INVENTOR(S): Bourbonnais, Yves, Cap-Rouge, Canada
Deslauriers, Noella, St-Ferreol-les-Neiges, Canada

PATENT ASSIGNEE(S): Universite Laval, Quebec, Canada (non-U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 6320033	B1	20011120
APPLICATION INFO.:	US 1999-351200		19990709 (9)

	NUMBER	DATE
PRIORITY INFORMATION:	CA 1998-2237134	19980710
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	GRANTED	
PRIMARY EXAMINER:	Marschel, Ardin H.	
LEGAL REPRESENTATIVE:	Nixon Peabody LLP	
NUMBER OF CLAIMS:	2	
EXEMPLARY CLAIM:	1	
NUMBER OF DRAWINGS:	19 Drawing Figure(s); 11 Drawing Page(s)	
LINE COUNT:	810	

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB The present invention relates to a Candida albicans gene (CSA1) encoding a surface protein. The present invention also relates to the surface protein and methods for using the protein or the gene for the detection, prophylaxis or treatment of candidal infection. The protein encoded is a surface antigen of Candida albicans yeast and mycelial forms, respectively. The mycelial surface antigen was shown to be present predominantly in the terminal third of the hyphal structures. CSA1 is a gene coding for a unique surface antigen.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

INCL INCLM: 536/023.100

INCLS: 435/006.000; 435/254.220; 536/023.740; 536/024.320

NCL NCLM: 536/023.100

NCLS: 435/006.000; 435/254.220; 536/023.740; 536/024.320

(FILE 'HCAPLUS, MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH, JICST-EPLUS, JAPIO, USPATFULL' ENTERED AT 09:47:30 ON 04 DEC 2003)

L68 10292 S "COX R"?/AU

Searcher : Shears 308-4994

10/081935

L69 1193 S "MAGEE D"?/AU
L70 5719 S "JIANG C"?/AU
L71 21 S L68 AND L69 AND L70
L72 97 S L68 AND (L69 OR L70)
L73 21 S L69 AND L70

-Author(s)

L75 93 SEA ABB=ON PLU=ON (L71 OR L72 OR L73 OR L68 OR L69 OR
L70) AND (L53 OR L30)
L76 25 DUP REM L75 (68 DUPLICATES REMOVED)

L76 ANSWER 1 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 1

ACCESSION NUMBER: 2003:790176 HCAPLUS

TITLE: Identification of a protective antigen of
Coccidioides immitis by expression
library immunization

AUTHOR(S): Ivey, F. Douglas; **Magee, D. Mitchell**;
Woitaske, Melanie D.; Johnston, Stephen Albert;
Cox, Rebecca A.

CORPORATE SOURCE: Department of Microbiology and Immunology,
University of Texas Health Science Center at San
Antonio, San Antonio, TX, 78229, USA

SOURCE: Vaccine (2003), 21(27-30), 4359-4367
CODEN: VACCDE; ISSN: 0264-410X

PUBLISHER: Elsevier Science Ltd.

DOCUMENT TYPE: Journal

LANGUAGE: English

AB **Coccidioides immitis** is a fungal pathogen of humans and is
classified as a Select Agent. We have identified a new potential
vaccine candidate for this pathogen using cDNA expression library
immunization (ELI). A **C. immitis** spherule-phase
cDNA library containing 800-1000 genes was divided into 10 pools and
each was tested for its protective capacity in BALB/c mice against
i.p. challenge with 2500 arthroconidia of this dimorphic fungus.
The most protective pool, designated Pool 7, was fractionated into
five sublibraries, each containing 60 genes, and of these, only Pool 7-3
induced a significant level of protection in mice. Fractionation of
Pool 7-3 into six sublibraries, each with 10 genes, yielded a
protective fraction, designated Pool 7-3-5. Subsequent fraction of
the latter pool into 10 sublibraries, each with one clone, yielded a
clone designated 7-3-5-5 that was highly protective. Clone 7-3-5-5
was sequenced and found to contain a 672 bp ORF encoding a 224 amino
acid **protein** having a 19 amino acid signal sequence on the
N-terminus and a 15 amino acid C-terminal GPI anchor site. The
7-3-5-5 clone, designated ELI-Antigen 1 (ELI-Ag1), showed partial
homol. with a hypothetical **protein** from *Neurospora crassa*.
This is the first study to identify a protective antigen from a
fungus using ELI, and it is also the first report in which
sequential fractionation of an expression library successfully
identified a single protective gene.

L76 ANSWER 2 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 2

ACCESSION NUMBER: 2003:721561 HCAPLUS

DOCUMENT NUMBER: 139:346526

TITLE: Transfection of murine dendritic cell line (JAWS
II) by a nonviral transfection reagent

AUTHOR(S): Awasthi, Shanjana; **Cox, Rebecca A.**

CORPORATE SOURCE: Dept. of Pathology, University of Texas Health

10/081935

Science Center at San Antonio, San Antonio, TX,
USA
SOURCE: BioTechniques (2003), 35(3), 600-602,604
CODEN: BTNQDO; ISSN: 0736-6205
PUBLISHER: Eaton Publishing Co.
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Dendritic cells are the most potent antigen-presenting cells that initiate and modulate the host immune system. Based on their immunostimulatory activity, a variety of strategies have been developed to use dendritic cells as vaccines and immunotherapeutic agents against infection and cancer. Genetically modified dendritic cells are useful for immunotherapeutic purposes because of their sustained activity in vivo. However, transfection of dendritic cells with plasmid DNA has been very difficult. While the viral transfection is associated with nonspecific activation of dendritic cells, commonly used nonviral transfection reagents have a low efficiency of transfection. Here we describe an improved, simple, less time-consuming transfection protocol using the nonviral nonliposomal lipid polymer, TransIT-TKO transfection reagent, for transfecting murine dendritic cells (JAWS II) with the gene that encodes **Coccidioides immitis** antigen 2 (Ag2). The JAWS II cells were cotransfected with pHYG-enhanced green fluorescent **protein** (EGFP) and pVR1012-**C immitis** Ag2 plasmid DNAs using TransIT-TKO reagent. We reproducibly obtained 30%-50% transfection efficiency. The transfected cells maintained their immature phenotype and were functionally active. In addition, the flexibility of this agent for expressing multiple antigens (GFP and **C. immitis** Ag2) offers an advantage of delivering multiple immunogens.

REFERENCE COUNT: 14 THERE ARE 14 CITED REFERENCES AVAILABLE
FOR THIS RECORD. ALL CITATIONS AVAILABLE
IN THE RE FORMAT

L76 ANSWER 3 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 3

ACCESSION NUMBER: 2002:489694 HCAPLUS

DOCUMENT NUMBER: 137:92359

TITLE: Role of signal sequence in vaccine-induced
protection against experimental
coccidioidomycosis

AUTHOR(S): Jiang, Chengyong; Magee, D.
Mitchell; Ivey, F. Douglas; Cox,
Rebecca A.

CORPORATE SOURCE: Department of Microbiology, University of Texas
Health Science Center at San Antonio, San
Antonio, TX, 78229, USA

SOURCE: Infection and Immunity (2002), 70(7), 3539-3545
CODEN: INFIBR; ISSN: 0019-9567

PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The vaccine efficacy of the gene sequence encoding the signal **peptide** of the antigen known as antigen 2 or proline-rich antigen (Ag2/PRA), an immunodominant antigen present in the cell wall of the fungal pathogen **Coccidioides immitis**, was investigated in a murine model of **coccidioidomycosis**. Expression plasmids for Ag2/PRA(1-18) DNA (signal sequence), Ag2/PRA(19-194) DNA (lacking the signal sequence), and

10/081935

Ag2/PRA(1-194) DNA (full length) were inserted in the pVR1012 vector, and the constructs were used to vaccinate the highly susceptible BALB/c mouse strain. Immunization with the signal gene sequence significantly reduced the fungal burden in the lungs and spleens of mice 12 days after i.p. challenge with a LD of 2500 *C. immitis* arthroconidia, to a level comparable to the protection induced in mice immunized with the full-length Ag2/PRA(1-194) DNA. The Ag2/PRA(19-194) gene protected mice but to a significantly lower level than the signal sequence or the full-length Ag2 gene. The immunizing capacity of Ag2/PRA(1-18) was not attributable to a nonspecific immunostimulatory effect of DNA, as evidenced by the fact that mice immunized with a frameshift mutation of Ag2/PRA(1-18) were not protected against challenge. Furthermore, a synthetic **peptide** corresponding to the translated sequence of Ag2/PRA(1-18) DNA protected mice, albeit at a lower level than the Ag2/PRA(1-18) DNA vaccine. The protection induced with the signal gene vaccine correlated with the production of gamma interferon when splenocytes from Ag2/PRA(1-18)-immunized mice were stimulated with recombinant full-length Ag2 and was not associated with the production of anti-*Coccidioides* IgG antibody. This is the first study to establish that a signal **peptide** sequence alone, administered as a gene vaccine or synthetic **peptide**, can induce protective immunity against a microbial pathogen.

REFERENCE COUNT: 36 THERE ARE 36 CITED REFERENCES AVAILABLE
FOR THIS RECORD. ALL CITATIONS AVAILABLE
IN THE RE FORMAT

L76 ANSWER 4 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC. on
STN

ACCESSION NUMBER: 2002:597076 BIOSIS

DOCUMENT NUMBER: PREV200200597076

TITLE: Induction of gene expression in *Coccidioides*
immitis during growth inside murine macrophages.

AUTHOR(S): **Magee, D. M.** [Reprint author]; **Woitaske, M.**
D. [Reprint author]; **Wickes, B. L.** [Reprint author];
Cox, R. A. [Reprint author]

CORPORATE SOURCE: University of Texas Health Science Center at San
Antonio, San Antonio, TX, USA

SOURCE: Abstracts of the General Meeting of the American
Society for Microbiology, (2002) Vol. 102, pp. 205.
print.
Meeting Info.: 102nd General Meeting of the American
Society for Microbiology. Salt Lake City, UT, USA.
May 19-23, 2002. American Society for Microbiology.
ISSN: 1060-2011.

DOCUMENT TYPE: Conference; (Meeting)
Conference; Abstract; (Meeting Abstract)

LANGUAGE: English

ENTRY DATE: Entered STN: 20 Nov 2002

Last Updated on STN: 20 Nov 2002

AB *Coccidioides immitis* is a dimorphic fungus of the American
southwest growing in the soil as mycelia and within the host as the
spherule/endospore parasitic phase. We hypothesized that genes that
are upregulated when *C. immitis* grows in vivo
might be associated with virulence of this primary fungal pathogen.
To identify potential virulence-associated genes, we utilized
differential display, reverse transcriptase, polymerase chain

10/081935

reaction (DD-RT-PCR) to assess induction of gene expression within *C. immitis* spherules during intracellular parasitism of murine macrophages. Macrophage monolayers were infected with *C. immitis* arthroconidia and, after approximately 36 hours of culture, RNA was obtained from the developing macrophage-derived spherules and compared to RNA from spherules grown in tissue culture medium alone. Differential display analyses identified 11 gene fragments specific for the macrophage-derived spherules. These products were cloned into a TA vector, sequenced, and analyzed for homology with the Genbank databases. One clone had significant homology to a *Magnaporthe grisea* appressorium stage cDNA library gene and *Saccharomyces cerevisiae* glycogen phosphorylase. The complete cDNA clone was recovered using 5' and 3' RACE and using this as a probe, we obtained and sequenced a genomic clone for *C. immitis* glycogen phosphorylase. The gene had seven predicted introns and a translation of the coding region indicated that the protein would contain 881 amino acids with a molecular weight of 100.6 kDa and a pI of 5.75. Northern analysis confirmed that glycogen phosphorylase message was differentially expressed by spherules in parasitized macrophages. Differential display may provide a tool to delineate virulence-associated genes of *C. immitis* or help define metabolic targets for rationally designed antifungal drugs.

L76 ANSWER 5 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 4

ACCESSION NUMBER: 2000:208899 HCAPLUS

DOCUMENT NUMBER: 132:247984

TITLE: The x-ray structure of a chitinase from the

AUTHOR(S): pathogenic fungus *Coccidioides immitis*
Hollis, Thomas; Monzingo, Arthur F.; Bortone,
Kara; Ernst, Stephen; Cox, Rebecca;
Robertus, Jon D.

CORPORATE SOURCE: Institute of Cellular and Molecular Biology,
Department of Chemistry and Biochemistry,
University of Texas, Austin, TX, 78712, USA

SOURCE: Protein Science (2000), 9(3), 544-551
CODEN: PRCLIE; ISSN: 0961-8368

PUBLISHER: Cambridge University Press

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The x-ray crystal structure of chitinase from *C.*

immitis was solved to 2.2 Å resolution. Like other members of the class 18 hydrolase family, this 427-residue protein was an 8-stranded β/α -barrel. Although lacking an N-terminal chitin-anchoring domain, the enzyme closely resembled chitinase from *Serratia marcescens*. Among the conserved features were 3 cis peptide bonds, all involving conserved active site residues. The active site was formed from conserved residues such as Trp-47, Trp-131, Trp-315, Trp-378, Tyr-239, Tyr-293, Arg-52 and Arg-295. Glu-171 was the catalytic acid in the hydrolytic mechanism; when it was mutated to Gln, enzyme activity was abolished. Allosamidin is a substrate analog that strongly inhibits the class 18 enzymes. Its binding to the chitinase, hevamine, was previously observed and conserved structural features of the 2 enzymes were used to predict the inhibitor's binding to the fungal enzyme.

REFERENCE COUNT: 39 THERE ARE 39 CITED REFERENCES AVAILABLE
FOR THIS RECORD. ALL CITATIONS AVAILABLE

10/081935

IN THE RE FORMAT

L76 ANSWER 6 OF 25 EMBASE COPYRIGHT 2003 ELSEVIER INC. ALL RIGHTS
RESERVED. on STN

ACCESSION NUMBER: 1999185649 EMBASE
TITLE: Construction of a single-chain interleukin-12-
expressing retroviral vector and its application in
cytokine gene therapy against experimental
coccidioidomycosis.

AUTHOR: Jiang C.; Magee D.M.; Cox
R.A.

CORPORATE SOURCE: C. Jiang, Dept. of Clinical Investigation, Texas
Center for Infectious Disease, 2303 SE Military Dr.,
San Antonio, TX 78223, United States.
chengyong.jiang@tdh.state.tx.us

SOURCE: Infection and Immunity, (1999) 67/6 (2996-3001).
Refs: 45
ISSN: 0019-9567 CODEN: INFIBR

COUNTRY: United States

DOCUMENT TYPE: Journal; Article

FILE SEGMENT: 004 Microbiology
026 Immunology, Serology and Transplantation

LANGUAGE: English

SUMMARY LANGUAGE: English

AB T-cell-mediated immunity is an important determinant in protection
against primary infection with **Coccidioides immitis**, a
dimorphic fungal pathogen that causes the disease
coccidioidomycosis. To determine if interleukin-12 (IL-12)
gene therapy could potentiate host response against **C.**
immitis, we constructed a single-chain cDNA encoding the p40
and p35 subunits linked by a polylinker and, using a retroviral
vector, transfected J774 macrophages with the construct. The
transduced J774 cells expressed IL-12 in vitro, with a mean
concentration of 28,440 pg from 106 cells in 48 h as measured by an
IL-12 (p75)-specific enzyme-linked immunosorbent assay. The secreted
IL-12 was biologically active, as judged by its ability to induce
the production of gamma interferon (IFN- γ) by spleen cells
from BALB/c mice. Treatment of the highly susceptible BALB/c mouse
strain with the IL-12- transduced J774 cells inhibited **C.**
immitis growth in tissues from mice challenged by a
pulmonary route, as evidenced by 1.37-, 2.59-, and 1.22-log
reductions in the number of CFU in the lungs, spleens, and livers,
respectively, compared to the fungal load in mice given
vector-transduced J774 cells. The protective effect of IL-12 gene
therapy was accompanied by increased levels of IFN- γ in the
lungs and sera of mice treated with IL-12- transduced J774 cells and
the constitutive production of IFN- γ by their spleen cells
cultured in vitro. These results suggest that IL-12 gene therapy
could be used as adjunct therapy for **coccidioidomycosis**.

L76 ANSWER 7 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 5

ACCESSION NUMBER: 1999:83280 HCAPLUS

DOCUMENT NUMBER: 130:266091

TITLE: Genetic vaccination against **Coccidioides**
immitis: comparison of vaccine efficacy of
recombinant antigen 2 and antigen 2 cDNA

AUTHOR(S): Jiang, Chengyong; Magee, D.
Mitchell; Quitugua, Teresa N.; Cox,

10/081935

CORPORATE SOURCE: Rebecca A.
Department of Clinical Investigation, Texas
Center for Infectious Disease, San Antonio, TX,
78223, USA

SOURCE: Infection and Immunity (1999), 67(2), 630-635
CODEN: INFIBR; ISSN: 0019-9567

PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Previous studies from our laboratory established that C-ASWS, an alkali-soluble, water-soluble extract from cell walls of *Coccidioides immitis*, protects mice against lethal challenge with this fungus. The C-ASWS extract contains a glycosylated **protein**, designated antigen 2 (Ag2), and a polysaccharide antigen. We recently cloned Ag2 cDNA and showed that the recombinant fusion **protein** elicited strong delayed-type hypersensitivity responses in immunized mice. This investigation was undertaken to determine if the recombinant Ag2 **protein**, expressed as an Ag2-glutathione S-transferase (GST) fusion **protein**, or Ag2 cDNA would protect mice against lethal challenge with *C. immitis*. The recombinant Ag2-GST **protein** protected BALB/c mice against i.p. challenge with 250 arthroconidia, as assessed by a decrease in fungal CFU in tissues. The Ag2-GST-immunized mice did not show, however, an increased survival during a 30-day period postinfection. By contrast, immunization of mice with Ag2 cDNA ligated into the pVR1012 plasmid engendered protection against i.p. challenge with 2,500 arthroconidia and against pulmonary challenge with 50 arthroconidia. Vaccine efficacy paralleled the development of delayed-type hypersensitivity responses to *C. immitis* antigen. Whereas mice vaccinated with the recombinant Ag2-GST **protein** did not mount footpad hypersensitivity to C-ASWS or the recombinant Ag2-GST **protein**, mice vaccinated with the pVR1012-Ag2 construct mounted a strong footpad hypersensitivity and their spleen cells secreted gamma interferon upon in vitro stimulation with the Ag2-containing C-ASWS extract. This is the first investigation to show that genetic immunization can protect against lethal challenge with *C. immitis*.

REFERENCE COUNT: 50 THERE ARE 50 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 8 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 6

ACCESSION NUMBER: 1999:2211 HCAPLUS

DOCUMENT NUMBER: 130:135866

TITLE: Crystallization and preliminary x-ray analysis of a chitinase from the fungal pathogen *Coccidioides immitis*

AUTHOR(S): Hollis, Thomas; Monzingo, Arthur F.; Bortone, Kara; Schelp, Elisabeth; Cox, Rebecca; Robertus, Jon D.

CORPORATE SOURCE: Institute of Cellular and Molecular Biology,
Department of Chemistry and Biochemistry,
University of Texas, Austin, TX, 78712, USA

SOURCE: Acta Crystallographica, Section D: Biological
Crystallography (1998), D54(6, Pt. 2), 1412-1413
CODEN: ABCRE6; ISSN: 0907-4449

PUBLISHER: Munksgaard International Publishers Ltd.

10/081935

DOCUMENT TYPE: Journal
LANGUAGE: English

AB Chitinase from *C. immitis* was expressed as a fusion **protein** with glutathione-S-transferase (GST), which aided in its purification. After cleavage from GST, the chitinase was crystallized from 30% PEG 4000 in 0.1M NaOAc, pH 4.6. The crystals had a tetragonal crystal lattice and belonged to space group P41212 or P43212 and diffracted to 2.2 Å resolution. The unit-cell parameters were $a = b = 91.2$, $c = 95.4$ Å; there is only 1 chitinase mol. in the asym. unit.

REFERENCE COUNT: 12 THERE ARE 12 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 9 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 7

ACCESSION NUMBER: 1997:655696 HCAPLUS

DOCUMENT NUMBER: 127:330024

TITLE: Mapping of a *Coccidioides immitis*-specific epitope that reacts with complement-fixing antibody

AUTHOR(S): Yang, Michael C.; Magee, D. Mitchell; Cox, Rebecca A.

CORPORATE SOURCE: Department of Clinical Investigation, Texas Center for Infectious Disease, San Antonio, TX, 78223, USA

SOURCE: Infection and Immunity (1997), 65(10), 4068-4074
CODEN: INFIBR; ISSN: 0019-9567

PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The authors have previously cloned the cDNA fragment that encodes the complement fixation antigen of *Coccidioides immitis*. The recombinant **protein** was highly sensitive in detecting CF antibody in sera from patients with **coccidioidomycosis** but was not specific to *C. immitis*, as evidence by its reactivity with sera from patients with histoplasmosis and, to lesser extent, blastomycosis. The authors undertook this study to determine if the epitope(s) that reacts with CF antibody is the same or differs from the epitopes that are shared with *Histoplasma capsulatum* and *Blastomyces dermatitidis*. PCR-generated CF/chitinase cDNA fragments were cloned and examined for their reactivity in enzyme-linked immunosorbent assays using sera from patients with **coccidioidomycosis**, histoplasmosis, or blastomycosis. A **peptide** domain comprised of amino acid residues 20 through 310 was shown to express an epitope(s) that is specific to anti-*Coccidioides* CF antibody. The **peptide** detected serum antibody in 21 (95%) of 22 patients with active **coccidioidomycosis** and was without reactivity with sera from 20 patients with histoplasmosis, 15 patients with blastomycosis, and 14 healthy subjects. Antibody titers to the recombinant **peptide** directly correlated with CF antibody titers ($P < 0.01$), and preadsorption of reference CF antiserum with the **peptide** ablated the reactivity of the antiserum in the immunodiffusion assay for CF antibody. The delineation of a recombinant **peptide** that has both sensitivity and specificity will provide a valuable tool for detecting CF antibody and for evaluating the role of CF antibody in the host response to *C. immitis*.

10/081935

L76 ANSWER 10 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 8
ACCESSION NUMBER: 1997:482804 HCAPLUS
DOCUMENT NUMBER: 127:204149
TITLE: Identification of a **Coccidioides**
immitis antigen 2 domain that expresses
B-cell-reactive epitopes
AUTHOR(S): Zhu, Yufan; Tryon, Vic; **Magee, D.**
Mitchell; Cox, Rebecca A.
CORPORATE SOURCE: Department of Clinical Investigation, Texas
Center for Infectious Disease, San Antonio, TX,
78223, USA
SOURCE: Infection and Immunity (1997), 65(8), 3376-3380
CODEN: INFIBR; ISSN: 0019-9567
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Antigen 2 (Ag2), a major immunoreactive component of **C. immitis** mycelium- and spherule-phase cell walls, was recently cloned and was shown to elicit T-cell responses in **Coccidioides**-immune mice. Here, the authors evaluated recombinant Ag2 (rAg2) and PCR-generated Ag2 truncations for expression of B-cell-reactive epitopes in ELISA and immunoblot assays with sera from patients with active **coccidioidomycosis**, a hyperimmune goat anti-Ag2 serum, and a murine anti-Ag2 monoclonal antibody that recognizes a conformational epitope. The results established that rAg2 expresses both linear and conformational B-cell-reactive epitopes which are localized to a domain comprised of amino acids 19-96 (designated A19-96). Truncations designed to identify epitopes within the A19-96 domain yielded fragments that either were nonreactive (A62-194, A19-61, and A49-79) or showed reduced reactivity (A19-79). Hence, A19-96 was the shortest domain expressing epitopes recognized by the panel of antibodies. The prevalence of antibodies to the A19-96 domain was evaluated in ELISAs of sera from 28 **coccidioidomycosis** patients. Antibody reactivity was detected in 79% of the patients' sera, and the level of antibody reactivity was directly correlated with disease severity. Whereas patients with pulmonary disease showed a mean response (A405) of 0.16, patients with disseminated **coccidioidomycosis** showed a mean response of 0.69. No reactivity was detected with sera from histoplasmosis or blastomycosis patients. The production of a recombinant **peptide** that expresses **C. immitis**-specific Ag2 epitopes provides a useful reagent for examining the role of anti-Ag2 antibodies in **coccidioidomycosis**.

L76 ANSWER 11 OF 25 MEDLINE on STN DUPLICATE 9
ACCESSION NUMBER: 97160980 MEDLINE
DOCUMENT NUMBER: 97160980 PubMed ID: 9008276
TITLE: Recombinant **Coccidioides immitis**
complement-fixing antigen: detection of an epitope
shared by **C. immitis**, *Histoplasma*
capsulatum, and *Blastomyces dermatitidis*.
AUTHOR: Yang M C; **Magee D M**; Kaufman L; Zhu Y;
Cox R A
CORPORATE SOURCE: Department of Clinical Investigation, Texas Center
for Infectious Disease, San Antonio 78223, USA.
CONTRACT NUMBER: AI12431 (NIAID)

Searcher : Shears 308-4994

10/081935

SOURCE: CLINICAL AND DIAGNOSTIC LABORATORY IMMUNOLOGY, (1997
Jan) 4 (1) 19-22.
Journal code: 9421292. ISSN: 1071-412X.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199704
ENTRY DATE: Entered STN: 19970422
Last Updated on STN: 19970422
Entered Medline: 19970408

AB We undertook an investigation to assess the utility of a recombinant **Coccidioides immitis** complement-fixing (CF) antigen for detecting CF antibody in sera from patients with **coccidioidomycosis**. Enzyme-linked immunosorbent assays established that recombinant CF antigen and, for comparison, a commercially available **coccidioidin** were reactive with 19 of 19 sera from patients with active **coccidioidomycosis**. The recombinant antigen was significantly more sensitive than **coccidioidin**. The median titer obtained when patients' sera were assayed against recombinant CF antigen was 1:51,200 compared to 1:25,600 with **coccidioidin** ($P < 0.027$). The recombinant antigen was also more effective in distinguishing the antibody levels in sera from patients with pulmonary **coccidioidomycosis** than in sera from those with disseminated disease. Whereas patients with pulmonary disease showed a median antibody titer of 1:25,600, those with multifocal disease showed a median titer of 1:102,400 ($P < 0.028$). The recombinant CF antigen was found, however, to express an epitope(s) that reacted with sera from 6 of 12 patients with histoplasmosis and 2 of 12 patients with blastomycosis.

L76 ANSWER 12 OF 25 MEDLINE on STN
ACCESSION NUMBER: 96355868 MEDLINE
DOCUMENT NUMBER: 96355868 PubMed ID: 8751906
TITLE: Interleukin-12 regulation of host defenses against **Coccidioides immitis**.
AUTHOR: Magee D M; Cox R A
CORPORATE SOURCE: Department of Clinical Investigation, Texas Center for Infectious Diseases, San Antonio 78223, USA..
mmagee@tcid.tdh.state.tx.us
CONTRACT NUMBER: AI23555 (NIAID)
SOURCE: INFECTION AND IMMUNITY, (1996 Sep) 64 (9) 3609-13.
Journal code: 0246127. ISSN: 0019-9567.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals; AIDS
ENTRY MONTH: 199610
ENTRY DATE: Entered STN: 19961015
Last Updated on STN: 19961015
Entered Medline: 19961003

AB We have previously reported on the alternate regulation of gamma interferon (IFN-gamma) and interleukin-4 (IL-4) in inbred mouse strains which differ in their susceptibility to **Coccidioides immitis**. The genetically resistant DBA/2 mice manifest a predominant T-helper 1 (Th1) response, with early production of IFN-gamma, whereas susceptible BALB/c mice show an early production

10/081935

of the Th2 cytokine IL-4. Since IL-12 is one cytokine that can act early during host defenses to promote the differentiation of cytokine production towards IFN-gamma and thus may promote expression of a protective immune response, we investigated the role of IL-12 in resistance to *C. immitis*.

Administration of recombinant IL-12 to the susceptible mouse strain before and after systemic (intraperitoneal) challenge with *C. immitis* significantly ameliorated the course of the disease, as measured by a reduction in the fungal load in the lungs, liver, and spleen. Analysis of the cytokine mRNA in lungs from infected BALB/c mice revealed that the protective effect of recombinant IL-12 was accompanied by a shift from a Th2 to a Th1 response. The importance of IL-12 in resistance to this fungus was further established by showing that neutralization of endogenous IL-12 in the resistant DBA/2 mouse strain led to a significant increase in the fungal burden in pulmonary and extrapulmonary tissues. These results establish that IL-12 plays a pivotal role in the host defense against systemic challenge with *C. immitis*.

L76 ANSWER 13 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 10
ACCESSION NUMBER: 1996:391996 HCAPLUS
DOCUMENT NUMBER: 125:55750
TITLE: Molecular cloning and characterization of
Coccidioides immitis antigen 2 cDNA
AUTHOR(S): Zhu, Yufan; Yang, Chunmu; Magee, D.
Mitchell; Cox, Rebecca A.
CORPORATE SOURCE: Dep. Clinical Investigation, Texas Center
Infectious Disease, San Antonio, TX, 78223, USA
SOURCE: Infection and Immunity (1996), 64(7), 2695-2699
CODEN: INFIBR; ISSN: 0019-9567
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Previous expts. have provided evidence that *Coccidioides immitis* antigen 2 (Ag2) is a major T-cell-reactive component of mycelia and spherule cell walls. Here the authors report the identification and cloning of the cDNA that encodes Ag2 from a lambda ZAP cDNA expression library constructed from spherule-derived RNA. DNA sequence anal. established that the 1,255-bp clone contains a 174-bp 5' untranslated region, a 582-bp open reading frame which encodes for a **protein** consisting of 194 amino acids, and a 375-bp 3' untranslated region, including a poly(A) tail. The recombinant Ag2 **protein** has a predicted mol. mass of 19.5 kDa and contains an 18-amino-acid N terminus which has been tentatively identified as a signal **peptide**. The Ag2 cDNA was ligated into the pGEX-4T-3 vector and expressed in *Escherichia coli* TG-1 cells as a glutathione S-transferase fusion **protein**. The recombinant fusion **protein** showed reactivity with sera from patients with **coccidioidomycosis** and elicited delayed-type footpad hypersensitivity responses in *Coccidioides*-immune mice. These results suggest that the Ag2 cDNA can be used for the large-scale production of this immunol. important **protein**.

L76 ANSWER 14 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 11
ACCESSION NUMBER: 1996:327965 HCAPLUS
DOCUMENT NUMBER: 125:28509

10/081935

TITLE: Molecular cloning and characterization of the **Coccidioides immitis** complement fixation/chitinase antigen

AUTHOR(S): Yang, Chunmu; Zhu, Yufan; Magee, D. Mitchell; Cox, Rebecca A.

CORPORATE SOURCE: Department Clinical Investigation, Texas Center Infectious disease, San Antonio, TX, 78223, USA

SOURCE: Infection and Immunity (1996), 64(6), 1992-1997
CODEN: INFIBR; ISSN: 0019-9567

PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Detection of anti-**Coccidioides** complement-fixing (CF) antibody is a valuable diagnostic and prognostic aid in **coccidioidomycosis**. The CF antibody response is directed against a heat-labile antigen that has chitinase activity, hereafter referred to as the CF/chitinase **protein**. To identify and clone this immunoreactive enzyme, the authors constructed a **Coccidioides immitis** cDNA lambda ZAP expression library from spherule RNA and detected fusion **peptides** expressing CF epitopes by immunoscreening. A cDNA clone consisting of 1623 bp was identified, sequenced, and found to contain a single open reading frame that encodes a **protein** of 47 kDa with 427 amino acids. Deduced amino acid sequence analyses showed that the cloned CF/chitinase cDNA contains a 35-amino-acid region, beginning at Ser-18 and ending at Arg-52, which has 92% homol. with the reported N-terminal amino acid sequence of authentic CF/chitinase **protein**. The first 17 amino acids in the deduced sequence of the cloned cDNA are not present on the mature CF/chitinase **protein**, suggesting that it may be a signal **peptide**. Expression of the CF/chitinase cDNA insert by using the pGEX-4T-3 vector yields a fusion **peptide** that bears CF-specific epitopes and shows chitinase activity. The CF/chitinase clone will enable large-scale production of the recombinant CF antigen for use in immunoassays and facilitate studies on the role of chitinase in the morphogenesis of **C immitis**.

L76 ANSWER 15 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 12

ACCESSION NUMBER: 1996:737195 HCAPLUS

DOCUMENT NUMBER: 126:86206

TITLE: **Coccidioides immitis** Antigen 2: analysis of gene and **protein**

AUTHOR(S): Zhu, Yufan; Yang, Chunmu; Magee, D. Mitchell; Cox, Rebecca A.

CORPORATE SOURCE: Department of Clinical Investigation, Texas Center for Infectious Disease, 2303 SE Military Drive, San Antonio, TX, 78223, USA

SOURCE: Gene (1996), 181(1/2), 121-125
CODEN: GENED6; ISSN: 0378-1119

PUBLISHER: Elsevier

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Antigen 2 is a glycosylated **protein** present in the cell walls of the dimorphic fungus **Coccidioides immitis**. Using oligodeoxyribonucleotide (oligo) primers based on the sequences of Ag2 cDNA, the gene encoding Ag2 was cloned from genomic DNA derived from the mycelial phase of **C. immitis** by PCR. Nucleotide (nt) sequence analyses showed a 582 base pair (bp) ORF

10/081935

disrupted by two introns which are 78bp and 101bp long. The deduced primary translation product consists of 194 amino acids (aa), contains an N-terminal putative signal sequence to allow transport into the endoplasmic reticulum, and a C-terminal putative signal sequence to enable a GPI anchor addition. Putative GPI anchor/cleavage site and O-glycosylation sites, as well as phosphorylation and myristoylation sites are also present. On the basis of these analyses, the authors predict that a prepro-Ag2 undergoes a post-translational modification to yield the mature glycosylated Ag2 **protein** which is anchored on the extracellular plasma membrane of mycelial and spherule-phase cells.

L76 ANSWER 16 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 13
ACCESSION NUMBER: 1995:781432 HCAPLUS
DOCUMENT NUMBER: 123:196430
TITLE: Roles of gamma interferon and interleukin-4 in
genetically determined resistance to

Coccidioides immitis
AUTHOR(S): Magee, D. Mitchell; Cox, Rebecca
A.

CORPORATE SOURCE: Dep. Res. Immunol., Texas Cent. Infectious
Diseases, San Antonio, TX, 78223, USA
SOURCE: Infection and Immunity (1995), 63(9), 3514-19
CODEN: INFIBR; ISSN: 0019-9567
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The profiles of γ interferon (IFN- γ) and interleukin-4 (IL-4) production were valuated during the course of **coccidioidomycosis** in 2 inbred mouse strains which differ in their susceptibility to **C. immitis**. Cytokine responses, measured at the mol. and **protein** levels, showed increased levels of IFN- γ in lung exts. from mice of the resistant DBA/2 strain after a pulmonary challenge, whereas the susceptible BALB/c strain manifested a predominant IL-4 response. The importance of these cytokines in host defense against **C. immitis** was established by treating the mice with recombinant cytokines or neutralizing anticytokine monoclonal antibodies. Treatment of the susceptible BALB/c mice with recombinant murine IFN- γ protected mice against systemic challenge, and in the reciprocal experiment, the administration of an anti-IFN- γ monoclonal antibody to the resistant DBA/2 mice decreased their capacity to control disease. Although the treatment of DBA/2 mice with recombinant IL-4 did not alter the disease, neutralization of endogenous IL-4 in infected BALB/c mice by administration of a neutralizing anti-IL-4 antibody led to a reduction in the fungal load in their tissues. Apparently, IFN- γ plays a pivotal role in resistance to **C. immitis**, whereas IL-4 down-regulates protective immunity against **C. immitis**.

L76 ANSWER 17 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 14
ACCESSION NUMBER: 1993:426295 HCAPLUS
DOCUMENT NUMBER: 119:26295
TITLE: Production of a murine monoclonal antibody that
recognizes an epitope specific to

Coccidioides immitis antigen 2
AUTHOR(S): Cox, Rebecca A.; Dolan, Matthew J.;

10/081935

CORPORATE SOURCE: Magee, D. Mitchell; Galgiani, John N.
Dep. Res. Immunol., San Antonio State Chest
Hosp., San Antonio, TX, 78223, USA
SOURCE: Infection and Immunity (1993), 61(5), 1895-9
CODEN: INFIBR; ISSN: 0019-9567
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Antigen 2 (Ag2) has been implicated as a T-cell-reactive component of the pathogenic fungus *C. immitis*. This report describes the production of a murine monoclonal antibody (MAb) of the IgG2a isotype that recognizes an epitope specific to *C. immitis* Ag2. This specificity was evidenced by the finding that the MAb did not recognize other antigens present in *coccidioidin* or spherulin and did not show reactivity with antigenic exts. from *Histoplasma capsulatum* or *Blastomyces dermatitidis*. The epitope was labile to enzymic digestion with pronase but resistant to treatment with glycolytic enzymes and to periodate oxidation. This **peptide** epitope appears to require conformational structure because it was not recognized by the MAb in immunoblots of antigen that had been electrophoresed in polyacrylamide gels under denaturing, reducing conditions. Immunoaffinity chromatog. of spherulin on columns containing the MAb established that the MAb was effective as a ligand for isolating Ag2 from heterogeneous exts. The production of a MAb which recognizes an Ag2-specific epitope and its utility as a ligand for isolating Ag2 will provide a valuable reagent for studies of this immunol. important antigen.

L76 ANSWER 18 OF 25 EMBASE COPYRIGHT 2003 ELSEVIER INC. ALL RIGHTS
RESERVED. on STN DUPLICATE 15

ACCESSION NUMBER: 92245388 EMBASE
DOCUMENT NUMBER: 1992245388
TITLE: Localization of the tube precipitin and complement fixation antigens of *Coccidioides immitis* by immunoelectron microscopy with murine monoclonal antibodies.
AUTHOR: Cox R.A.; Sun S.H.; Dolan M.J.; Harrison J.L.
CORPORATE SOURCE: Department of Research Immunology, San Antonio State Chest Hospital, San Antonio, TX 78223, United States
SOURCE: Infection and Immunity, (1992) 60/8 (3315-3324).
ISSN: 0019-9567 CODEN: INFIBR
COUNTRY: United States
DOCUMENT TYPE: Journal; Article
FILE SEGMENT: 004 Microbiology
005 General Pathology and Pathological Anatomy
026 Immunology, Serology and Transplantation
LANGUAGE: English
SUMMARY LANGUAGE: English
AB The cellular localization of the tube precipitin (TP) and complement fixation (CF) antigens of *Coccidioides immitis* was examined by immunoelectron microscopy with murine immunoglobulin G1 monoclonal antibodies directed against the TP and CF antigens, respectively. Immunoelectron microscopic analyses of saprobic- and parasitic-phase cells showed that the TP antigen is present at a high concentration within the inner cell wall layer and along the plasma membrane. The antigen was also detected, at a lesser concentration, within cytoplasmic vacuoles. In contrast to the

10/081935

predominant localization of the TP antigen in the cell walls, the CF antigen resides primarily within the cytoplasm, where it appears to be dispersed throughout the cytoplasm rather than associated with a specific cytoplasmic organelle. A sparse amount of the CF antigen within the inner cell walls was also demonstrable. The localization of the TP and CF antigens throughout the morphogenetic phases of **C. immitis** has important implications in antigen production and in analyses of host response in **coccidioidomycosis**.

L76 ANSWER 19 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 16
ACCESSION NUMBER: 1991:533582 HCAPLUS
DOCUMENT NUMBER: 115:133582
TITLE: An immunoreactive apoglycoprotein purified from **Coccidioides immitis**
AUTHOR(S): Dugger, Kris O.; Galgiani, John N.; Ampel, Neil M.; Sun, Sung H.; **Magee, D. Mitchell**; Harrison, Jeff; Law, John H.
CORPORATE SOURCE: Med. Res. Serv., Veterans Aff. Med. Cent., Tucson, AZ, 85723, USA
SOURCE: Infection and Immunity (1991), 59(7), 2245-51
CODEN: INFIBR; ISSN: 0019-9567
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Deglycosylation of glycoproteins in a lysate of spherules of **C. immitis** has permitted purification and partial characterization of a proline-rich pronase-sensitive antigen. Moreover, soluble antigen specifically stimulated lymphocytes from persons with dermal delayed-type hypersensitivity to **coccidioidal** antigens. When related to reference **coccidioidin** by tandem 2-dimensional immunoelectrophoresis, the antigen fused in the anodal region with a specific reference antigen (antigen 2). It did not show identity with **coccidioidal** antigens used in conventional serol. assays. Although immunoblots of the purified **protein** with monospecific rabbit antiserum showed a single antigen at 33 kDa, the parent spherule lysate bound the same antibody in a broad band between 70 and >200 kDa, which could be explained by microheterogeneity of glycosylation. Immunoelectron microscopy using affinity-purified human antibodies localized the antigen to the cell wall and internal septa of spherules. The apoglycoprotein may be important in human immune responses to **coccidioidal** infection.

L76 ANSWER 20 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 17
ACCESSION NUMBER: 1991:490408 HCAPLUS
DOCUMENT NUMBER: 115:90408
TITLE: Production and characterization of a monoclonal antibody to the complement fixation antigen of **Coccidioides immitis**
AUTHOR(S): Dolan, Matthew J.; **Cox, Rebecca A.**
CORPORATE SOURCE: Dep. Res. Immunol., San Antonio State Chest Hosp., San Antonio, TX, 78223, USA
SOURCE: Infection and Immunity (1991), 59(6), 2175-80
CODEN: INFIBR; ISSN: 0019-9567
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Detection of complement-fixing antibody to **coccidioidin** by using the complement fixation test or an immunodiffusion assay for

10/081935

complement-fixing antibody (IDCF) is widely viewed as the most useful immunodiagnostic test for **coccidioidomycosis**. Here is reported the production of an IgG1 monoclonal antibody (MAB) to the IDCF antigen for use as a biospecific ligand for purifying the IDCF antigen on solid-phase immunosorbents and for use as a reagent for screening genomic or cDNA expression libraries from **C. immitis**. BALB/c mice were immunized by i.m. injections of **coccidioidin** in adjuvant, followed by an intrasplenic booster injection of **coccidioidin** in saline. The spleen cells were fused with SP2/0 Ag14 myeloma cells, and the fusion products were screened for IgG antibody to **coccidioidin** by using an ELISA. Pos. hybridomas were cloned and evaluated for reactivity to the IDCF antigen by 2-dimensional immunoelectrophoresis and by immunoblotting. The epitope recognized by the MAB was heat labile (60°, 30 min) and susceptible to enzymic digestion with pronase but was resistant to treatment with lipase, α -mannosidase, glucose oxidase, and endoglycosidase H. This heat-labile **peptide** epitope appears to be specific to **C. immitis**, as judged by the fact that the MAB was not reactive in immunoblots or ELISAs of histoplasmin or blastomycin.

L76 ANSWER 21 OF 25 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 18

ACCESSION NUMBER: 1989:455564 HCAPLUS

DOCUMENT NUMBER: 111:55564

TITLE: Induction of tumor necrosis factor alpha by spherules of **Coccidioides immitis**

AUTHOR(S): Slagle, David C.; Cox, Rebecca A.; Kuruganti, Uma

CORPORATE SOURCE: Dep. Res. Immunol., San Antonio State Chest Hosp., San Antonio, TX, 78223, USA

SOURCE: Infection and Immunity (1989), 57(7), 1916-21
CODEN: INFIBR; ISSN: 0019-9567

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The cytokine tumor necrosis factor α (TNF- α) functions as an immunomodulatory **protein** and as a mediator of cachexia. Here, the viable or formalin-killed spherules of **C. immitis** induced the secretion of TNF- α by peritoneal-exudate cells from BALB/c mice. The identification of the cytokine as TNF- α was based on its lytic activity against the TNF- α -sensitive Ls murine fibrosarcoma cell line but not the TNF- α -resistant LR cell line, its neutralization by rabbit anti-TNF- α , and its secretion by peritoneal cells having characteristics of macrophages. The induction of TNF- α was attributable to spherules and not to contaminating lipopolysaccharide (endotoxin), as evidenced by the finding that polymyxin B, a reagent that blocks the TNF- α -inducing component of lipopolysaccharide, did not negate the production of TNF- α in response to spherules, whereas pretreatment of spherules with hyperimmune goat antiserum to spherulin neutralized the induction of TNF- α by these cells. The demonstration that **C. immitis** activates macrophages to secrete TNF- α in vitro is a new finding and warrants studies to determine whether this cytokine is produced during active **coccidioidomycosis**.

L76 ANSWER 22 OF 25 MEDLINE on STN

DUPLICATE 19

Searcher : Shears 308-4994

10/081935

ACCESSION NUMBER: 88085464 MEDLINE
DOCUMENT NUMBER: 88085464 PubMed ID: 3335400
TITLE: Induction and expression of cell-mediated immune responses in inbred mice infected with **Coccidioides immitis**.
AUTHOR: Cox R A; Kennell W; Boncyk L; Murphy J W
CORPORATE SOURCE: Department of Research Immunology, San Antonio State Chest Hospital, Texas 78223.
CONTRACT NUMBER: AI23555 (NIAID)
SOURCE: INFECTION AND IMMUNITY, (1988 Jan) 56 (1) 13-7.
Journal code: 0246127. ISSN: 0019-9567.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 198802
ENTRY DATE: Entered STN: 19900305
Last Updated on STN: 19970203
Entered Medline: 19880209

AB Comparisons of the course of **coccidioidomycosis** in two strains of inbred mice established that BALB/c mice are significantly more susceptible to pulmonary infection with **Coccidioides immitis** than are DBA/2 mice. The susceptibility of BALB/c mice does not reside in their inability to mount a delayed-type hypersensitivity response to **C. immitis** antigen. That is, BALB/c mice manifested footpad hypersensitivity to **coccidioidin** early during the course of disease, to a level comparable to that of DBA/2 mice. In contrast to the more resistant DBA/2 mouse strain, however, BALB/c mice developed anergy by day 15 postinfection. Suppression of the delayed-type hypersensitivity response was not specific for **C. immitis** antigen, as evidenced by the finding that BALB/c mice immunized with mycobacterial purified **protein** derivative prior to infection with **C. immitis** were suppressed in their footpad response to mycobacterial antigen at day 15 postinfection. Taken together, these results establish that genetically determined susceptibility to this fungus is associated with an acquired suppression of cell-mediated immune reactivity.

L76 ANSWER 23 OF 25 MEDLINE on STN DUPLICATE 20
ACCESSION NUMBER: 87193089 MEDLINE
DOCUMENT NUMBER: 87193089 PubMed ID: 3552984
TITLE: Serum-mediated suppression of lymphocyte transformation responses in **coccidioidomycosis**.
AUTHOR: Cox R A; Pope R M
CONTRACT NUMBER: AI 18761 (NIAID)
SOURCE: INFECTION AND IMMUNITY, (1987 May) 55 (5) 1058-62.
Journal code: 0246127. ISSN: 0019-9567.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 198706
ENTRY DATE: Entered STN: 19900303
Last Updated on STN: 19970203
Entered Medline: 19870605

10/081935

AB Lymphocyte transformation (LT) responses to **coccidioidin** (CDN) and spherulin were suppressed in 11 (73%) of 15 patients with active **coccidioidomycosis** when their mononuclear cells were assayed in autologous serum as compared to serum from healthy, CDN skin test-positive subjects. Suppressed LT responses were specific for **Coccidioides immitis** antigens in 7 (64%) of the 11 patients. Immunoaffinity chromatography of patient sera with *Staphylococcus* **protein A** adsorbed the suppressor component(s) and thereby established that suppression was attributed to immunoglobulin G, either alone or complexed with antigen. The possibility that suppression was mediated by immune complexes was examined by adding complexes formed in vivo or in vitro to mononuclear cell cultures of healthy CDN-reactive persons before LT assays. Although complexes prepared in this manner were reactive in an enzyme-linked immunosorbent assay designed to detect **Coccidioides** antigen-specific immune complexes, no suppression of LT responses was observed. We conclude that serum-mediated suppression of LT responses in **coccidioidomycosis** is attributed to monomeric and not immune-complexed immunoglobulin G antibody.

L76 ANSWER 24 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC. on STN

ACCESSION NUMBER: 1984:33723 BIOSIS
DOCUMENT NUMBER: PREV198426033723; BR26:33723
TITLE: CHEMICAL COMPOSITION OF A MYCELIAL AND SPHERULE CELL WALL ANTIGEN FROM **COCCIDIODES-IMMITIS**.
AUTHOR(S): OLSBERG C A [Reprint author]; **COX R A**
CORPORATE SOURCE: UNIV TEX HEALTH SCI CENT, SAN ANTONIO, TX, USA
SOURCE: Abstracts of the Annual Meeting of the American Society for Microbiology, (1983) Vol. 83, pp. F6. Meeting Info.: 83RD ANNUAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, NEW ORLEANS, LA., USA, MAR. 6-11, 1983. ABSTR ANNU MEET AM SOC MICROBIOL. CODEN: ASMACK. ISSN: 0094-8519.
DOCUMENT TYPE: Conference; (Meeting)
FILE SEGMENT: BR
LANGUAGE: ENGLISH

L76 ANSWER 25 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC. on STN

ACCESSION NUMBER: 1978:79432 BIOSIS
DOCUMENT NUMBER: PREV197815022932; BR15:22932
TITLE: CHEMICAL COMPOSITION OF AN ALKALI SOLUBLE WATER SOLUBLE ANTIGEN FROM **COCCIDIODES-IMMITIS**.
AUTHOR(S): FLEMING W H III; LAND G A; **COX R A**
SOURCE: Abstracts of the Annual Meeting of the American Society for Microbiology, (1978) Vol. 78.1978, pp. 314. CODEN: ASMACK. ISSN: 0094-8519.
DOCUMENT TYPE: Article
FILE SEGMENT: BR
LANGUAGE: Unavailable

FILE 'HOME' ENTERED AT 09:51:03 ON 04 DEC 2003

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 13, 2003, 12:34:43 ; Search time 214 Seconds
(without alignments)
274.896 Million cell updates/sec

Title: US-10-081-935-2

Perfect score: 82

Sequence: 1 MFSSHAIALVAAGLASA 18

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10081935/runat_12112003_144408_3740/app.query.fasta_1.199
-DB=Published Applications NA -QFMT=fastap -SUPFI=rnpb -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=40 -MODES=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10081935.@CGN_1.1_164.@runat_12112003_144408_3740
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FCGAPOP=6 -FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	46	56.1	168	12	US-10-029-386-16850, A

C	2	46	56.1	569	12	US-10-029-386-3150	Sequence 3150, Ap
C 3	44	53.7	548	12	US-10-027-632-95893	Sequence 95893, A	
C 4	44	53.7	548	12	US-10-027-632-305915	Sequence 305915, A	
C 5	44	53.7	548	13	US-10-027-632-95893	Sequence 95893, A	
C 6	44	53.7	548	13	US-10-027-632-305915	Sequence 305915, A	
C 7	44	53.7	9372	12	US-10-293-582-2	Sequence 2, Appli	
C 8	43	52.4	549	11	US-09-991-936-889	Sequence 889, App	
C 9	43	52.4	549	11	US-09-991-936-889	Sequence 889, App	
C 10	43	52.4	720	10	US-09-894-844-66	Sequence 1638, Ap	
C 11	43	52.4	11115	12	US-09-769-734-49	Sequence 66, Appl	
C 12	42	51.2	2632	13	US-10-108-605-36	Sequence 49, Appl	
C 13	42	51.2	3042	10	US-09-712-363-68	Sequence 36, Appl	
C 14	41	50.0	251	14	US-10-198-846-12310	Sequence 12310, A	
C 15	41	50.0	498	11	US-09-991-936-975	Sequence 975, App	
C 16	41	50.0	609	10	US-09-738-626-3337	Sequence 3337, Ap	
C 17	41	50.0	894	14	US-10-198-846-4458	Sequence 4458, Ap	
C 18	41	50.0	1218	14	US-10-137-473-3	Sequence 3, Appli	
C 19	41	50.0	1722	14	US-10-156-761-4387	Sequence 4387, Ap	
C 20	41	50.0	2123	11	US-09-866-050A-584	Sequence 584, App	
C 21	41	50.0	2123	14	US-10-152-661-584	Sequence 584, App	
C 22	41	50.0	2484	14	US-10-156-761-918	Sequence 918, Appl	
C 23	41	50.0	3309400	10	US-09-738-626-1	Sequence 1, Appli	
C 24	41	50.0	9025608	14	US-10-156-761-1	Sequence 5360, Ap	
C 25	40	48.8	228	9	US-09-294-093B-5360	Sequence 70, Appl	
C 26	40	48.8	309	12	US-10-190-305A-71	Sequence 71, Appl	
C 27	40	48.8	309	12	US-10-190-305A-71	Sequence 66199, A	
C 28	40	48.8	406	12	US-10-027-632-66199	Sequence 66199, A	
C 29	40	48.8	406	13	US-10-027-632-66199	Sequence 4623, Ap	
C 30	40	48.8	436	10	US-09-867-701-4623	Sequence 36340, A	
C 31	40	48.8	445	11	US-09-918-995-36266	Sequence 36340, A	
C 32	40	48.8	467	12	US-10-027-632-36340	Sequence 36340, A	
C 33	40	48.8	467	13	US-10-027-632-36340	Sequence 42872, A	
C 34	40	48.8	509	13	US-10-027-632-42872	Sequence 42872, A	
C 35	40	48.8	509	13	US-10-027-632-42872	Sequence 281505, A	
C 36	40	48.8	512	12	US-10-027-632-281505	Sequence 281505, A	
C 37	40	48.8	512	13	US-10-027-632-281505	Sequence 30945, A	
C 38	40	48.8	513	11	US-09-918-995-30945	Sequence 30945, A	
C 39	40	48.8	622	12	US-10-027-632-185077	Sequence 185077, A	
C 40	40	48.8	622	12	US-10-027-632-185077	Sequence 185077, A	
C 41	40	48.8	622	12	US-10-027-632-185077	Sequence 185077, A	
C 42	40	48.8	622	13	US-10-027-632-185077	Sequence 185077, A	
C 43	40	48.8	622	13	US-10-027-632-185077	Sequence 185077, A	
C 44	40	48.8	622	13	US-10-027-632-185077	Sequence 203917, A	
C 45	40	48.8	637	12	US-10-027-632-203917	Sequence 203917, A	

ALIGNMENTS

RESULT 1

US-10-029-386-16850/c
; Sequence 16850, Application US/10029386
; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

; FILE REFERENCE: AEMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 16850

; LENGTH: 168

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO CHR.X.1

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.9

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1

```
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
; OTHER INFORMATION: SWISSPROT HIT: Q93008, EVALUE 5.00e-20
; OTHER INFORMATION: NT HIT: gi14764091, EVALUE 9.00e-71
; OTHER INFORMATION: EST_HUMAN HIT: AW996853.1, EVALUE 5.00e-73
US-10-029-386-16850
Alignment Scores:
Pred. No.: 2.37 Length: 168
Score: 46.00 Matches: 9
Percent Similarity: 76.47% Conservative: 4
Best Local Similarity: 52.94% Mismatches: 4
Query Match: 56.10% Indels: 0
DB: 12 Gaps: 0

US-10-081-935-2 (1-18) x US-10-029-386-16850 (1-168)
Qy 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSer 17
Db 110 CTTCAACACAATCATGCCCTAGTTACTTTGGTAGCAGAAAACCTTGCAACT 60
RESULT 2
US-10-029-386-3150/c
; Sequence 3150, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Shatton G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 2001-12-20
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3150
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR.X.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
; OTHER INFORMATION: SWISSPROT HIT: Q93008, EVALUE 2.00e-19
; OTHER INFORMATION: EST_HUMAN HIT: AW996853.1, EVALUE 5.00e-79
; OTHER INFORMATION: NT HIT: gi14764084, EVALUE 3.00e-70
US-10-029-386-3150
Alignment Scores:
Pred. No.: 10.8 Length: 569
Score: 46.00 Matches: 9
Percent Similarity: 76.47% Conservative: 4
Best Local Similarity: 52.94% Mismatches: 4
Query Match: 56.10% Indels: 0
DB: 12 Gaps: 0

US-10-081-935-2 (1-18) x US-10-029-386-3150 (1-569)
Qy 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSer 17
Db 276 CTTCAACACAATCATGCCCTAGTTACTTTGGTAGCAGAAAACCTTGCAACT 226
RESULT 3
US-10-027-632-95893
; Sequence 95893, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95893
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-95893
Alignment Scores:
Pred. No.: 25.6 Length: 548
Score: 44.00 Matches: 8
Percent Similarity: 83.33% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 53.66% Indels: 0
DB: 12 Gaps: 0

US-10-081-935-2 (1-18) x US-10-027-632-95893 (1-548)
Qy 3 PheSerHisAlaLeuIleAlaLeuValAlaAlaGly 14
Db 320 TTTTCCCATATGATGATAGCCTCATCACTACTGCTGGG 355
RESULT 4
US-10-027-632-305915
; Sequence 305915, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 305915
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305915
```

Alignment Scores:
Pred. No.: 25.6 Length: 548
Score: 44.00 Matches: 8
Percent Similarity: 83.33% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 53.66% Indels: 0
DB: 12 Gaps: 0

US-10-081-935-2 (1-18) x US-10-027-632-305915 (1-548)

Qy 3 PheSerHisAlaLeuIleAlaLeuValAlaAlaGly 14
||||| :|||||
Db 320 TTTTCCCATATGATAGCACTCATAACTGCTGGG 355

RESULT 5
US-10-027-632-95893
; Sequence 95893, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 95893
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-95893

Alignment Scores:
Pred. No.: 25.6 Length: 548
Score: 44.00 Matches: 8
Percent Similarity: 83.33% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 53.66% Indels: 0
DB: 12 Gaps: 0

US-10-081-935-2 (1-18) x US-10-027-632-95893 (1-548)

Qy 3 PheSerHisAlaLeuIleAlaLeuValAlaAlaGly 14
||||| :|||||
Db 320 TTTTCCCATATGATAGCACTCATAACTGCTGGG 355

RESULT 6
US-10-027-632-305915
; Sequence 305915, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 305915
LENGTH: 548
TYPE: DNA
ORGANISM: Human
US-10-027-632-305915

Alignment Scores:
Pred. No.: 25.6 Length: 548
Score: 44.00 Matches: 8
Percent Similarity: 83.33% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 53.66% Indels: 0
DB: 12 Gaps: 0

US-10-081-935-2 (1-18) x US-10-027-632-305915 (1-548)

Qy 3 PheSerHisAlaLeuIleAlaLeuValAlaAlaGly 14
||||| :|||||
Db 320 TTTTCCCATATGATAGCACTCATAACTGCTGGG 355

RESULT 7
US-10-293-582-2
; Sequence 2, Application US/10293582
; Publication No. US20030175253A1
; GENERAL INFORMATION:
; APPLICANT: Akil, Huda
; APPLICANT: Bunney, William E.
; APPLICANT: Burke, Sharon
; APPLICANT: Choudary, Prabhakara V.
; APPLICANT: Cox, David R.
; APPLICANT: Evans, Simon
; APPLICANT: Jones, Edward G.
; APPLICANT: Li, Jun
; APPLICANT: Lopez, Juan F.
; APPLICANT: The Trustees of The Leland Stanford Junior University
; TITLE OF INVENTION: Compositions and Methods for Diagnosing and Treating
; FILE REFERENCE: 020885-0002100S
; CURRENT APPLICATION NUMBER: US/10/293,582
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/339,252
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 2
; LENGTH: 9372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 9 (USP9Y) (DFFRY)
US-10-293-582-2

Alignment Scores:
Pred. No.: 881 Length: 9372
Score: 44.00 Matches: 9
Percent Similarity: 76.47% Conservative: 4
Best Local Similarity: 52.94% Mismatches: 4
Query Match: 53.66% Indels: 0

Alignment Scores:
Pred. No.: 40.4 Length: 549
Score: 43.00 Matches: 9
Percent Similarity: 72.22% Conservative: 4
Best Local Similarity: 50.00% Mismatches: 5
Query Match: 52.44% Indels: 0
DB: 11 Gaps: 0

US-10-081-935-2 (1-18) x US-09-991-936-1638 (1-549)

QY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
|||:::||:::||:::||:::||
Db 21 ATGAATAATTCAGTAGCAATTTTGGCCCTGGCCCTGTGTGGCATCAGCT 174

RESULT 10
US-09-894-844-66
; Sequence 66, Application US/09994844
; Patent No. US20020176873A1
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; TITLE OF INVENTION: the M. Tuberculosis Complex
; FILE REFERENCE: STAN10XCON
; CURRENT APPLICATION NUMBER: US/09/894,844
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 720
; TYPE: DNA
; ORGANISM: M. tuberculosis
US-09-894-844-66

Alignment Scores:
Pred. No.: 56.6 Length: 720
Score: 43.00 Matches: 10
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 90.91% Mismatches: 0
Query Match: 52.44% Indels: 0
DB: 10 Gaps: 0

US-10-081-935-2 (1-18) x US-09-894-844-66 (1-720)

QY 8 IleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
|||||:::|||:::|||:::|||
Db 25 ATCGCCCTGTGGCCGCCGCCTTGGCCCGCA 57

RESULT 11
US-09-769-734-49/C
; Sequence 49, Application US/09769734
; Publication No. US20030143666A1
; GENERAL INFORMATION:
; APPLICANT: Ecopia BioSciences Inc.
; TITLE OF INVENTION: Genetic Locus for Everninomicin Biosynthesis
; FILE REFERENCE: PA 005-US
; CURRENT APPLICATION NUMBER: US/09/769,734
; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 49
; LENGTH: 11115
; TYPE: DNA
; ORGANISM: M. carbonacea
; FEATURE:
; NAME/KEY: misc feature

US-10-198-846-12310/c
; Sequence 12310, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James

```
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12310
; LENGTH: 251
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12310

Alignment Scores:
Pred. No.: 37.6 Length: 251
Score: 41.00 Matches: 9
Percent Similarity: 71.43% Conservative: 1
Best Local Similarity: 64.23% Mismatches: 4
Query Match: 50.00% Indels: 0
DB: 14 Gaps: 0

US-10-081-935-2 (1-18) x US-10-198-846-12310 (1-251)

Qy 5 HisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
Db 122 CACGCGATGATCGTCTTTCGCCGCCGACGCTGTGTCGGCT 81

RESULT 15
US-09-991-936-975/c
; Sequence 975, Application US/09991936
; Publication No. US20030073827A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 975
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-975

Alignment Scores:
Pred. No.: 88.4 Length: 498
Score: 41.00 Matches: 11
Percent Similarity: 68.75% Conservative: 0
Best Local Similarity: 68.75% Mismatches: 3
Query Match: 50.00% Indels: 2
DB: 11 Gaps: 1

US-10-081-935-2 (1-18) x US-09-991-936-975 (1-498)
```

```
Qy 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAla 16
Db 445 ATGGATTTCCTCAC-----ATAACGTTAGTTGCAAAAGGATTAGCT 404

RESULT 16
US-09-738-626-3337
; Sequence 3337, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3337
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3337

Alignment Scores:
Pred. No.: 114 Length: 609
Score: 41.00 Matches: 5
Percent Similarity: 76.47% Conservative: 8
Best Local Similarity: 47.06% Mismatches: 4
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

US-10-081-935-2 (1-18) x US-09-738-626-3337 (1-609)

Qy 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSer 17
Db 400 CTTGATCGGATAGCGCCATCATCGTATCCGCCGACGAGGTCTCGCGCT 450

RESULT 17
US-10-198-846-4458/c
; Sequence 4458, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4458
; LENGTH: 894
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7, 15, 16, 22, 24, 27, 33, 37, 40, 80, 81, 85, 474, 484,
; LOCATION: 501, 512, 527, 543, 550, 551, 553, 555, 564, 567, 584, 587,
; LOCATION: 591, 592, 617, 618, 619, 623, 633, 634, 640, 642, 643, 651,
; LOCATION: 655, 659, 660, 664, 666, 667, 673, 674, 683, 684, 686
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 688, 689, 693, 695, 699, 700, 703, 713, 719, 724, 726, 727,
; LOCATION: 728, 729, 731, 734, 745, 746, 748, 751, 752, 753, 756, 757,
; LOCATION: 760, 762, 764, 766, 773, 780, 782, 783, 792, 796, 799, 800,
; LOCATION: 812, 813, 814, 823, 826, 839, 847, 848, 857, 861, 869
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 870, 872, 874, 881, 882, 888, 889, 890, 891
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-4458

Alignment Scores:
Pred. No.: 183 Length: 894
Score: 41.00 Matches: 9
Percent Similarity: 71.43% Conservative: 1
Best Local Similarity: 64.29% Mismatches: 4
Query Match: 50.00% Indels: 0
DB: 14 Gaps: 0

US-10-081-935-2 (1-18) x US-10-198-846-4458 (1-894)

Qy 5 HisAlaLeuValAlaLeuValAlaAlaGlyLeuAlaSerAla 18
Db 167 CACGGCATGATGCTCTTGGCCGCCACCGCTGCTGGCT 126

RESULT 18
US-10-137-473-3
; Sequence 3, Application US/10137473
; Publication No. US20030104419A1
; GENERAL INFORMATION:
; APPLICANT: Gould-Rothberg
; APPLICANT: Rastelli
; TITLE OF INVENTION: METHOD OF CLASSIFYING A THYROID CARCINOMA USING
; FILE REFERENCE: 15966-548
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US/10137,473
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/130,123
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 60/193,203
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Murinae gen. sp.
US-10-137-473-3

Alignment Scores:
Pred. No.: 269 Length: 1218
Score: 41.00 Matches: 8
Percent Similarity: 66.67% Conservative: 4
Best Local Similarity: 44.44% Mismatches: 6
Query Match: 50.00% Indels: 0
DB: 14 Gaps: 0

US-10-081-935-2 (1-18) x US-10-137-473-3 (1-1218)

Qy 1 MetGlnPheSerHisAlaLeuValAlaAlaGlyLeuAlaSerAla 18
Pred. No.: 1 Length: 1218
Score: 41.00 Matches: 8
```

```
Db 480 CTGCCTTTTGAGCCACAGCTCGTCTTCTCTGGCAATGGGTCTGGTCACAGCA 533

RESULT 19
US-10-156-761-4387/c
; Sequence 4387, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 4387
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1722)
US-10-156-761-4387

Alignment Scores:
Pred. No.: 415 Length: 1722
Score: 41.00 Matches: 9
Percent Similarity: 86.67% Conservative: 4
Best Local Similarity: 60.00% Mismatches: 2
Query Match: 50.00% Indels: 0
DB: 14 Gaps: 0

US-10-081-935-2 (1-18) x US-10-156-761-4387 (1-1722)

Qy 1 MetGlnPheSerHisAlaLeuValAlaAlaGlyLeu 15
Db 492 GTCCAGGACGCGCGTGCAGTTCCTCCACTGTGTCGCGGGCCTT 448

RESULT 20
US-09-866-050A-584
; Sequence 584, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 584
; LENGTH: 2123
; TYPE: DNA
; ORGANISM: Mouse
US-09-866-050A-584

Alignment Scores:
Pred. No.: 539 Length: 2123
Score: 41.00 Matches: 8
```

Percent Similarity: 66.67% Conservatives: 4
Best Local Similarity: 44.44% Mismatches: 6
Query Match: 50.00% Indels: 0
DB: 11 Gaps: 0

US-10-081-935-2 (1-18) x US-09-866-050A-584 (1-2123)

QY 1 MetGlnPheSerHisAlaLeuValAlaAlaGlyLeuAlaSerAla 18
DB 870 CTGCCTTTGAGCCACGCTGCTTCTCTCTGGAATGGTCTGTCACAGCA 923

RESULT 21

US-10-152-661-584
; Sequence 584, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c5
; CURRENT APPLICATION NUMBER: US/10/152,661
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 584
; LENGTH: 2123
; TYPE: DNA
; ORGANISM: Mouse
US-10-152-661-584

Alignment Scores:
Pred. No.: 539 Length: 2123
Score: 41.00 Matches: 8
Percent Similarity: 66.67% Conservatives: 4
Best Local Similarity: 44.44% Mismatches: 6
Query Match: 50.00% Indels: 0
DB: 14 Gaps: 0

US-10-081-935-2 (1-18) x US-10-152-661-584 (1-2123)

QY 1 MetGlnPheSerHisAlaLeuValAlaAlaGlyLeuAlaSerAla 18
DB 870 CTGCCTTTGAGCCACGCTGCTTCTCTCTGGAATGGTCTGTCACAGCA 923

RESULT 22

US-10-156-761-918
; Sequence 918, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI

Alignment Scores:
Pred. No.: 3e+06 Length: 3309400
Score: 41.00 Matches: 8
Percent Similarity: 76.47% Conservatives: 5
Best Local Similarity: 47.06% Mismatches: 4
Query Match: 50.00% Indels: 0

APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 918
LENGTH: 2484
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2484)
US-10-156-761-918

Alignment Scores:
Pred. No.: 655 Length: 2484
Score: 41.00 Matches: 8
Percent Similarity: 70.59% Conservatives: 4
Best Local Similarity: 47.06% Mismatches: 5
Query Match: 50.00% Indels: 0
DB: 14 Gaps: 0

US-10-081-935-2 (1-18) x US-10-156-761-918 (1-2484)

QY 1 MetGlnPheSerHisAlaLeuValAlaAlaGlyLeuAlaSer 17
DB 1168 CTGCAACTGACACACCGCTTGTCTGCTGAGCGGGCTGCGCCGCC 1218

RESULT 23
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Alignment Scores:
Pred. No.: 3e+06 Length: 3309400
Score: 41.00 Matches: 8
Percent Similarity: 76.47% Conservatives: 5
Best Local Similarity: 47.06% Mismatches: 4
Query Match: 50.00% Indels: 0

```
DB: 10 Gaps: 0
US-10-081-935-2 (1-18) x US-09-738-626-1 (1-3309400)
Qy 1 MetGlnPheSerHisAlaLeuValAlaLeuValAlaGlyLeuAlaSer 17
Db 3218392 CTTGACATCGGATACGCCATCATCGGTATCGCGCAGCAGGTCTCGCGCT 3218442
RESULT 24
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1
Alignment Scores:
Pred. No.: 4.25e+06 Length: 9025608
Score: 41.00 Matches: 8
Percent Similarity: 70.59% Conservative: 4
Best Local Similarity: 47.06% Mismatches: 5
Query Match: 50.00% Indels: 0
DB: 14 Gaps: 0
US-10-081-935-2 (1-18) x US-10-156-761-1 (1-9025608)
Qy 1 MetGlnPheSerHisAlaLeuValAlaLeuValAlaGlyLeuAlaSer 17
Db 1126427 CTGCAACTGACACACCCCTTGCTGCACCTTCGTCGAGCGGGGCTGGCGCC 1126477
RESULT 25
US-09-294-093B-5360/c
; Sequence 5360, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Laigudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 5360
; LENGTH: 228
; TYPE: DNA
; ORGANISM: Zea mays
```

```
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700356446H1
US-09-294-093B-5360
Alignment Scores:
Pred. No.: 52.5 Length: 228
Score: 40.00 Matches: 8
Percent Similarity: 84.62% Conservative: 3
Best Local Similarity: 61.54% Mismatches: 2
Query Match: 48.78% Indels: 0
DB: 9 Gaps: 0
US-10-081-935-2 (1-18) x US-09-294-093B-5360 (1-228)
'y 1 MetGlnPheSerHisAlaLeuValAlaLeuValAlaGlyLeuAlaSer 13
Db 153 ATCAATTGAATCATGCCCTTCTGCTCTTGTCAAAGCT 115
RESULT 26
US-10-190-305A-70/c
; Sequence 70, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: TatC22Exon1_2_TV1_C_ZAopt
US-10-190-305A-70
Alignment Scores:
Pred. No.: 76.7 Length: 309
Score: 40.00 Matches: 8
Percent Similarity: 78.57% Conservative: 3
Best Local Similarity: 57.14% Mismatches: 3
Query Match: 48.78% Indels: 0
DB: 12 Gaps: 0
US-10-081-935-2 (1-18) x US-10-190-305A-70 (1-309)
Qy 5 HisAlaLeuValAlaLeuValAlaGlyLeuAlaSerAla 18
Db 273 CACCTTCTTCTGCTCTCTCGCTGCGGTGCTGCTGCCCGG 232
RESULT 27
US-10-190-305A-71/c
; Sequence 71, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
```

```

US-10-027-632-66199/c
; Sequence 66199, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66199
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(406)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-66199

Alignment Scores:
Pred. No.: 108 Length: 406
Score: 40.00 Matches: 9
Percent Similarity: 75.00% Conservative: 3
Best Local Similarity: 56.25% Mismatches: 4
Query Match: 48.78% Indels: 0
DB: 13 Gaps: 0

US-10-081-935-2 (1-18) x US-10-027-632-66199 (1-406)

Qy 3 PheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 211 TTTTCGAATGCTCTATGAGCTTTTACTTGGGGGGTGCATTGCTCAGCT 164

RESULT 30
US-09-867-701-4623
; Sequence 4623, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4623
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-4623

Alignment Scores:
Pred. No.: 118 Length: 436
Score: 40.00 Matches: 8

```

```

; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36340
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(467)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-36340

Alignment Scores:
Pred. No.: 128                      Length: 467
Score: 40.00                      Matches: 9
Percent Similarity: 75.00%        Conservative: 3
Best Local Similarity: 56.25%     Mismatches: 4
Query Match: 48.78%              Indels: 0
DB: 12                           Gaps: 0

US-10-081-935-2 (1-18) x US-10-027-632-36340 (1-467)

Qy 3 PheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla i8
      |||||:|||||:|||||:|||||:
Db 211 TTTTCGAATGCTCTATGCTTTACTTGGGGTGGTTCATTGCTCACGT 164

RESULT 33
US-10-027-632-36340/c
; Sequence 36340, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36340
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(467)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-36340

Alignment Scores:
Pred. No.: 128                      Length: 467
Score: 40.00                      Matches: 9
Percent Similarity: 75.00%        Conservative: 3
Best Local Similarity: 56.25%     Mismatches: 4
Query Match: 48.78%              Indels: 0
DB: 13                           Gaps: 0

US-10-081-935-2 (1-18) x US-10-027-632-36340 (1-467)

```



```
Oy      3 PheSerHisAlaLeuValAlaAlaGlyLeuAlaSerAla 18
|||||:|||||:|||||:|||||:|||||
Db      211 TTTTGAATGCTCTTATGGCTTTACTTGGGGTGGTCTGCTCAGCT 164
|||||:|||||:|||||:|||||:|||||

RESULT 34
US-10-027-632-42872
; Sequence 42872, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42872
; LENGTH: 509
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-42872

Alignment Scores:
Pred. No.:      143      Length:      509
Score:          40.00    Matches:      7
Percent Similarity: 90.91% Conservative: 3
Best Local Similarity: 63.64% Mismatches: 1
Query Match:    48.78% Indels:      0
DB:             12      Gaps:        0

US-10-081-935-2 (1-18) x US-10-027-632-42872 (1-509)
Oy      3 PheSerHisAlaLeuValAlaAla 13
|||||:|||||:|||||:|||||:|||||
Db      369 TTTTCTCATTCCTTACTTGCATAGTAGGGGCC 401
|||||:|||||:|||||:|||||:|||||

RESULT 35
US-10-027-632-42872
; Sequence 42872, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
```

```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42872
; LENGTH: 509
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-42872

Alignment Scores:
Pred. No.:      143      Length:      509
Score:          40.00    Matches:      7
Percent Similarity: 90.91% Conservative: 3
Best Local Similarity: 63.64% Mismatches: 1
Query Match:    48.78% Indels:      0
DB:             13      Gaps:        0

US-10-081-935-2 (1-18) x US-10-027-632-42872 (1-509)
Oy      3 PheSerHisAlaLeuValAlaAla 13
|||||:|||||:|||||:|||||:|||||
Db      369 TTTTCTCATTCCTTACTTGCATAGTAGGGGCC 401
|||||:|||||:|||||:|||||:|||||

RESULT 36
US-10-027-632-281505
; Sequence 281505, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281505
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281505

Alignment Scores:
Pred. No.:      144      Length:      512
Score:          40.00    Matches:      7
Percent Similarity: 80.00% Conservative: 5
Best Local Similarity: 46.67% Mismatches: 3
Query Match:    48.78% Indels:      0
DB:             12      Gaps:        0

US-10-081-935-2 (1-18) x US-10-027-632-281505 (1-512)
Oy      4 SerHisAlaLeuValAlaAlaGlyLeuAlaSerAla 18
|||||:|||||:|||||:|||||:|||||
Db      112 TCCCATGCACTAGTGCATGATCATTTATTCGCGCTCTCCAGT 156
|||||:|||||:|||||:|||||:|||||

RESULT 37
```

```
US-10-027-632-281505
; Sequence 281505, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 10827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281505
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281505

Alignment Scores:
Pred. No.: 144      Length: 512
Score: 40.00      Matches: 7
Percent Similarity: 80.00%      Conservative: 5
Best Local Similarity: 46.67%      Mismatches: 3
Query Match: 48.78%      Indels: 0
DB: 13      Gaps: 0

US-10-081-935-2 (1-18) x US-10-027-632-281505 (1-512)
Oy 4 SerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
Db 112 TCCCATGCACGTAGTGGCCATGATCATTTATTGGCAGCTCCTCCACT 156

RESULT 38
US-09-918-995-30945
; Sequence 30945, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30945
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(513)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30945

Alignment Scores:
Pred. No.: 144      Length: 513
Score: 40.00      Matches: 8
```

```
Percent Similarity: 71.43%      Conservative: 2
Best Local Similarity: 57.14%      Mismatches: 4
Query Match: 48.78%      Indels: 0
DB: 11      Gaps: 0

US-10-081-935-2 (1-18) x US-09-918-995-30945 (1-513)
Oy 4 SerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSer 17
Db 231 AGCCATGCTCTGCTCACCATGGCCGGTGTGCTGCACCTGCCTCC 272

RESULT 39
US-10-027-632-185077
; Sequence 185077, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 10827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 185077
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-185077

Alignment Scores:
Pred. No.: 183      Length: 622
Score: 40.00      Matches: 8
Percent Similarity: 83.33%      Conservative: 2
Best Local Similarity: 66.67%      Mismatches: 2
Query Match: 48.78%      Indels: 0
DB: 12      Gaps: 0

US-10-081-935-2 (1-18) x US-10-027-632-185077 (1-622)
Oy 3 PheSerHisAlaLeuIleAlaLeuValAlaAlaGly 14
Db 436 TTCTCCCACTCCTGCTAGTATGAGGAGCCCAATGCT 471

RESULT 40
US-10-027-632-185078
; Sequence 185078, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 10827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
```

```

; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 185078
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-185078

```

```

Alignment Scores:
Pred. No.: 183 Length: 622
Score: 40.00 Matches: 8
Percent Similarity: 83.33% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 48.78% Indels: 0
DB: 12 Gaps: 0

```

US-10-081-935-2 (1-18) x US-10-027-632-185078 (1-622)

```

Qy 3 PheSerHisAlaLeuIleAlaLeuValAlaLaGly 14
   |||||:::|||||
Db 436 TTCTCCCACTCACTGGTAGCATTGGAGGCCAATGGT 471

```

Search completed: November 13, 2003, 14:12:16
Job time : 984 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 21:23:48 ; Search time 242 Seconds
(without alignments)
580.045 Million cell updates/sec

Title: US-10-081-935-1
Perfect score: 52
Sequence: 1 atgcagttcttcacgctct.....cgctgcggcctcgcagcg 52

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*			
1:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*		
2:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*		
3:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*		
4:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*		
5:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*		
6:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*		
7:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*		
8:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*		
9:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*		
10:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*		
11:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*		
12:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*		
13:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*		
14:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*		
15:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*		
16:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*		
17:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*		
18:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*		
19:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*		
20:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*		
21:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*		
22:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*		
23:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*		
24:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*		
25:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	25.6	49.2	1006	23 AAS59682 Propionibacterium
2	25.6	49.2	34088	23 AAS59566 Propionibacterium
3	25	48.1	336	22 AAD20424 Pseudomonas syring
4	25	48.1	1920	23 AAS54321 Pseudomonas aerugi
5	24.8	47.7	2016	23 ABL07659 Drosophila melanog
6	24.8	47.7	4095	23 ABL07659 Drosophila melanog
7	24.6	47.3	1566	22 AAS57171 DNA encoding Dros
8	24.6	47.3	1566	22 AAH49426 D. melanogaster pe

C 9	24.6	47.3	1566	23 ABL22999	Drosophila melanog
C 10	24.6	47.3	3772	22 AAS57170	DNA encoding Dros
C 11	24.6	47.3	3772	23 ABL22998	Drosophila melanog
C 12	24.4	46.9	695	21 AAF12291	Aspergillus oryzae
C 13	24.4	46.9	5484	23 AAS73006	DNA encoding novel
C 14	24.4	46.9	25913	22 AAK70862	Human immune/haema
C 15	24.4	46.9	198161	24 ABK83564	Human cDNA differe
C 16	24	46.2	2162	21 AAA94724	Rice auxin transpo
C 17	23.8	45.8	609	22 AAH68302	C glutamicum codin
C 18	23.8	45.8	732	22 AAF71300	Corynebacterium gl
C 19	23.8	45.8	7077	25 ABZ59094	E. maxima 230 kDa
C 20	23.8	45.8	7077	25 ABZ22977	E. maxima 250
C 21	23.8	45.8	7987	25 ABZ22976	E. maxima 250
C 22	23.8	45.8	7990	25 ABZ22984	E. maxima immunodo
C 23	23.8	45.8	49377	19 AAV05287	The soraphen biosy
C 24	23.8	45.8	58857	21 AAS48471	Nucleotide sequenc
C 25	23.8	45.8	309400	22 AAH68534	C glutamicum codin
C 26	23.6	45.4	1415	24 ABZ11723	Human polynucleoti
C 27	23.6	45.4	135638	25 ABX34289	S. atroolivaceus 1
C 28	23.4	45.0	145	24 AAD38151	pCON3 deletion mut
C 29	23.4	45.0	450	24 ABN20845	Human ORFX polynuc
C 30	23.4	45.0	855	21 AA94737	Corn auxin transpo
C 31	23.4	45.0	2466	21 AAC55847	Mitomycin biosynth
C 32	23.4	45.0	2840	23 ABL13923	Drosophila melanog
C 33	23.4	45.0	5736	23 ABL08443	Drosophila melanog
C 34	23.4	45.0	6782	23 ABL13922	Drosophila melanog
C 35	23.4	45.0	9577	23 ABL08442	Drosophila melanog
C 36	23.4	45.0	18034	21 AAC55841	Complete Mitomycin
C 37	23.2	44.6	271	25 ABX30608	Human GDP-mannose
C 38	23.2	44.6	450	24 ABQ98305	Human ORF112 codin
C 39	23.2	44.6	867	23 ABL26959	Drosophila melanog
C 40	23.2	44.6	1563	24 ABL60268	Trametes versicolo
C 41	23.2	44.6	1899	16 AAT05530	Metalloprotease co
C 42	23.2	44.6	2052	16 AAT05529	p45 metalloproteas
C 43	23.2	44.6	2052	17 AAT40133	Fusarium oxysporum
C 44	23.2	44.6	2052	19 AAV28868	Fusarium oxysporum
C 45	23.2	44.6	5437	12 AAQ13369	Fusaric acid resis

ALIGNMENTS

RESULT 1
AAS59682
ID AAS59682 standard; DNA; 1006 BP.

XX AAS59682;

AC AAS59682;

DT 13-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein encoding DNA #177.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant; ds.

XX Propionibacterium acnes.

OS Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

PA Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;


```
Query Match      47.3%; Score 24.6; DB 22; Length 1566;
Best Local Similarity 70.2%; Pred. No. 59;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      3 GCAGTTCTCTCAGGCTCTCATCGCTCTCGTCTCGTCCGCGCTCGCCA 49
DB      646 GCACAGCGGTACGACCTCTCGCTCTCTCGTCTCGTGGAGTCTCCACA 600

RESULT 8
AAH49426/c
ID   AAH49426 standard; cDNA; 1566 BP.
XX
AC   AAH49426;
XX
DT   11-DEC-2001 (first entry)
XX
DE   D. melanogaster peptide receptor encoding cDNA SEQ ID 43.
XX
KW   Insect; fruitfly; peptide receptor; plant protection; insecticide; ss.
XX
OS   Drosophila melanogaster.
XX
PN   DE10013618-Al.
XX
PD   20-SEP-2001.
XX
PF   18-MAR-2000; 2000DE-1013618.
XX
PR   18-MAR-2000; 2000DE-1013618.
XX
PA   (FARB ) BAYER AG.
XX
PI   Antonicek H, Friedrich G, Schulte T;
XX
WPI; 2001-571695/65.
DR   P-PSDB; AAB86973.
XX
New polypeptides from Drosophila melanogaster have biological activity
PT of peptide receptor, useful to find new compounds for plant protection
PT and insecticides -
XX
PS   Claim 6; Page 107-110; 128pp; German.
XX
This invention describes novel polypeptides (p1) from Drosophila
CC melanogaster having the biological activity of a peptide receptor.
CC Molecules of the invention are used to find new plant protection
CC compounds or insecticides, or to find genes encoding a polypeptide
CC involved in the structure of functionally similar receptors in insects
CC This sequence encodes a Drosophila melanogaster (fruitfly) peptide
CC receptor described in the method of the invention.
XX
SQ   Sequence 1566 BP; 291 A; 490 C; 437 G; 348 T; 0 other;

Query Match      47.3%; Score 24.6; DB 22; Length 1566;
Best Local Similarity 70.2%; Pred. No. 59;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      3 GCAGTTCTCTCAGGCTCTCATCGCTCTCGTCTCGTCCGCGCTCGCCA 49
DB      646 GCACAGCGGTACGACCTCTCGCTCTCTCGTCTCGTGGAGTCTCCACA 600

RESULT 9
ABL22999/c
ID   ABL22999 standard; DNA; 1566 BP.
XX
AC   ABL22999;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 20470.
```

```
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 20470; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511); expressed DNA
CC sequences (ABJ01840-ABJ16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1566 BP; 291 A; 490 C; 437 G; 348 T; 0 other;

Query Match      47.3%; Score 24.6; DB 23; Length 1566;
Best Local Similarity 70.2%; Pred. No. 59;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      3 GCAGTTCTCTCAGGCTCTCATCGCTCTCGTCTCGTCCGCGCTCGCCA 49
DB      646 GCACAGCGGTACGACCTCTCGCTCTCTCGTCTCGTGGAGTCTCCACA 600

RESULT 10
AAS57170/c
ID   AAS57170 standard; DNA; 3772 BP.
XX
AC   AAS57170;
XX
DT   16-JAN-2002 (first entry)
XX
DE DNA encoding Drosophila G-protein coupled receptor, GPCR #50.
XX
KW Drosophila; G-protein coupled receptor; GPCR; insecticide; diagnostic;
KW mutation detection; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200170980-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09341.
XX
PR 23-MAR-2000; 2000US-191638P.
PR 18-JUL-2000; 2000US-0618893.
```


PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-02559678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-483426/52.
DR
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 25674; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 25913 BP; 5398 A; 6705 C; 7089 G; 6721 T; 0 other;

Query Match 46.9%; Score 24.4; DB 22; Length 25913;
Best Local Similarity 68.0%; Pred. No. 96;
Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 2 TGCAGTTCTCTCACGCTCTCATCGCTCGCTGCGCTCGCGCTCGCCAGT 51
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1099 TGCAGTTCCCTCACAGTCTTCACGTTCTCTTCCTTCGGAAGCTCCAGT 1148

RESULT 15
 ABK83564/c
 ID ABK83564 standard; cDNA; 198161 BP.
 AC ABK83564;
 XX
 DT 14-AUG-2002 (first entry)
 XX
 DE Human cDNA differentially expressed in granulocytic cells #135.
 XX
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 XX
 OS Homo sapiens.
 XX
 PN WO200228999-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 03-OCT-2001; 2001WO-US30821.
 XX
 PR 03-OCT-2000; 2000US-237189P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 XX
 DR WPI; 2002-435328/46.
 XX
 PT Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity -
 XX
 PS Claim 1; SEQ ID No 135; 114pp; English.
 XX
 CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing
 CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gs is indicative of GCA.
 CC Also included are modulating (M2) GA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease using the
 CC gene expression profile; (3) detecting (M4) an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease, by detecting the
 CC level of expression in a sample of the tissue of gene(s) from Gs, where
 CC the level of expression of the gene is indicative of inflammation;
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 CC an allergic response in a subject, exposure of a subject to a pathogen
 CC or sterile inflammatory disease, by contacting a tissue having
 CC inflammation with an agent that modulates the expression of gene(s)
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
 CC modulating GA; M3 is useful for screening an agent capable of modulating
 CC GCA preferably in an inflammation in a tissue; M4 is useful for
 CC detecting an inflammation (especially chronic) in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC periodontal disease; also bacterial infection, viral infection,
 CC parasitic infection, protozoal infection, fungal infection and M5 is

CC useful for treating one of the above conditions. The present
 CC sequence represents a gene differentially expressed in granulocytes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 198161 BP; 61561 A; 45946 C; 41360 G; 49294 T; 0 other;
 XX
 Query Match 46.9%; Score 24.4; DB 24; Length 198161;
 Best Local Similarity 68.0%; Pred. No. 1.2e+02;
 Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 XX
 QY 2 TGCAGTTCCTCAGCGTCTCATCGCTCTGCTGCTGCGCGCTCGCCAGT 51
 Db 25296 TGCAGTTCCTCAGCGTCTCATCGTCTCTTCCTTCGCGAAAGCTCCAGT 25247
 XX
 RESULT 16
 AAA94724
 ID AAA94724 standard; DNA; 2162 BP.
 XX
 AC AAA94724;
 XX
 DT 02-FEB-2001 (first entry)
 XX
 DE Rice auxin transport protein clone rsl1n.pk003.n3 DNA sequence.
 XX
 KW Auxin transport protein; rice;
 KW root development; gene mapping; plant breeding; herbicide; ss.
 XX
 OS Oryza sativa.
 XX
 PN WO200068389-A2.
 XX
 PD 16-NOV-2000.
 XX
 PF 03-MAY-2000; 2000WO-US12061.
 XX
 PR 07-MAY-1999; 99US-0133040.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;
 XX
 DR WPI; 2000-687647/67.
 DR P-PSDB; AAB26937.
 XX
 PT New nucleic acid sequences encoding new auxin transport proteins,
 PT useful for modulating root growth of plants and to screen for
 PT herbicides -
 XX
 PS Claim 2; Pages 65-66; 94pp; English.
 XX
 CC Auxins are plant hormones that influence plant behaviour and development
 CC e.g. vascular tissue differentiation, apical development, tropic
 CC responses and organ (e.g. flower, leaf) development. The present
 CC invention relates to corn auxin transport protein coding sequences and
 CC proteins. The present sequence is one such coding sequence. This sequence
 CC may be used to modulate root development, e.g. to produce a more robust
 CC root system, alter root angle or redirect root growth. Also, the present
 CC sequence may be useful for gene mapping (e.g. for plant breeding) and to
 CC identify loss of function mutants. The protein encoded by the present
 CC sequence may be useful for raising specific antibodies, for the detection
 CC of auxin transport proteins and to design and/or identify specific
 CC inhibitors of auxin transport proteins, potentially useful as herbicides.
 XX
 SQ Sequence 2162 BP; 428 A; 694 C; 636 G; 404 T; 0 other;
 XX
 Query Match 46.2%; Score 24; DB 21; Length 2162;
 Best Local Similarity 68.8%; Pred. No. 97;
 Matches 33; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

AAAF71300

CC cysteine, valine, leucine, isoleucine, arginine, proline, histidine,
CC tyrosine, phenylalanine, or tryptophan. The fine chemical production can
CC be modulated. The presence of (1) or HA proteins encoded by then are
CC used for diagnosing the presence or activity of *Corynebacterium*
CC diphtheriae. (1) can be used to map the C. glutamicum genome or can be
CC used as markers for genetically engineered *Corynebacterium* or
CC *Brevibacterium*. The HA proteins encoded by the (1) are used to maintain
CC homeostasis in C. glutamicum or help the microorganism to adapt to
CC different environmental conditions.
XX
SQ Sequence 732 BP; 183 A; 220 C; 166 G; 163 T; 0 other;

Query Match 45.8%; Score 23.8; DB 22; Length 732;
Best Local Similarity 66.7%; Pred. No. 1.e+02;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 2 TGCAAGTCTCTCAGGCTCTCATCGCTCTCGTCTCGGCGGCTCGCAGTG 52
DB 501 TTCAGATCGGATACGCCCATCATCGGTATCCGCGCAGCAGGTCTGCGCGTG 551

RESULT 19
AB259094/C
ID AB259094 standard; cDNA; 7077 BP.

XX AC AB259094;
XX
XX 28-APR-2003 (first entry)
XX
XX E. maxima 230 kDa gametocyte antigen cDNA.

XX Gametocyte; antigen; 56 kDa; 82 kDa; antiparasitic; vaccine;
KW 230 kDa; Gene; ss.

XX Eimeria maxima.
XX
PN WO2003004683-A2.

XX 16-JAN-2003.

XX 03-JUL-2002; 2002WO-US21233.

XX 06-JUL-2001; 2001US-303699P.

XX (BELL/) BELLI S I.

XX (SMIT/) SMITH N C.

XX (WALL/) WALLACH M.

XX Belli SI, Smith NC, Wallach M;

XX WPI; 2003-221605/21.

XX New nucleic acid comprising a sequence encoding a 56 kDa or 82 kDa
PT polypeptide from gametocytes of *Eimeria maxima*, useful for preparing a
PT vaccine against *Eimeria* infection -

XX Disclosure; Fig 12; 112pp; English.

XX The invention provides two major E. maxima gametocyte antigens having
CC molecular weights of 56 and 82 kDa and encoding polynucleotides. The
CC gametocyte antigens can be produced by standard recombinant methodology.
CC The encoding polynucleotides are useful for preparing a vaccine against
CC *Eimeria* tenella, *E. acervulina*, *E. necatrix*, *E. praecox*, *E. maxima*,
CC *E. mitis* or *E. brunetti* infection. The present sequence represents a
CC E. maxima 230 kDa gametocyte antigen cDNA clone.
XX

SQ Sequence 7077 BP; 1814 A; 1987 C; 2163 G; 1113 T; 0 other;

Query Match 45.8%; Score 23.8; DB 25; Length 7077;
Best Local Similarity 72.1%; Pred. No. 1.3e+02;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 7 TTCTCTCAGGCTCTCATCGCTCTCGTCTCGGCGGCTCGGCCA 49

DB 2781 TTCTTCTGCGCACTCATCAATGCTTGGCAGCGCTCGCTCGCCA 2739

RESULT 20

AB222977/C

ID AB222977 standard; DNA; 7077 BP.

XX AC AB222977;

XX 10-APR-2003 (first entry)

XX Eimeria maxima 250 kDa protein homologous DNA sequence SEQ ID NO:4.
XX Eimeria maxima; 250 kDa antigen; sporozoite; merozoite; infection;
KW vaccine; antiparasitic; gene therapy; immunodominant; gene; ds.
XX Eimeria maxima.
OS Synthetic.

XX WO2003004684-A2.

XX 16-JAN-2003.

XX 03-JUL-2002; 2002WO-US21237.

XX 06-JUL-2001; 2001US-303670P.

XX (WITC/) WITCOMBE D.

XX (SMIT/) SMITH N C.

XX (WALL/) WALLACH M.

XX Witcombe D, Smith NC, Wallach M;

XX WPI; 2003-201556/19.

XX New nucleic acid comprising a sequence encoding a 250 kDa polypeptide
PT from sporozoites/merozoites of *Eimeria maxima*, useful for preparing a
PT vaccine against *Eimeria* infection -
XX Claim 15; Fig 7A; 198pp; English.

XX The present invention describes a nucleic acid (1) comprising a
CC sequence encoding a 250 kDa polypeptide from Sporozoites/Merozoites of
CC Eimeria maxima or its homologue or complement. Also described: (1) a
CC vector comprising the nucleic acid; (2) a host cell comprising the
CC vector; (3) a plasmid comprising the nucleic acid; (4) a transformed
CC cell comprising the nucleic acid; (5) producing a recombinant 250 kDa
CC polypeptide; (6) a recombinant polypeptide; (7) a vaccine against
CC E. tenella, E. acervulina, E. necatrix, E. praecox, E. maxima, E. mitis
CC or E. brunetti or a microorganism expressing an immunologically
CC cross-reactive antigen; (8) immunising a subject against infection by
CC Eimeria, or a microorganism expressing an immunologically cross-reactive
CC antigen; (9) conferring upon a newborn subject of an avian species
CC maternal immunity against infection by Eimeria; (10) a fertilised egg
CC from an avian species having an air sac, where the air sac is inoculated
CC with the vaccine; and (11) reducing the output of Eimeria oocytes in
CC faeces from a newborn subject of an avian species. (1) has antiparasitic
CC activity and can be used in a vaccine and in gene therapy. The nucleic
CC acid is useful for preparing a vaccine against E. tenella, E. acervulina,
CC E. necatrix, E. praecox, E. maxima, E. mitis or E. brunetti infection.
CC The present sequence represents a 250 kDa antigen protein homologous DNA
CC sequence, which is used in an example from the present invention.

SQ Sequence 7077 BP; 1814 A; 1985 C; 2165 G; 1113 T; 0 other;

Query Match 45.8%; Score 23.8; DB 25; Length 7077;
Best Local Similarity 72.1%; Pred. No. 1.3e+02;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 7 TTCTCTCAGGCTCTCATCGCTCTCGTCTCGGCGGCTCGGCCA 49

DB 2781 TTCTTCTGCGCACTCATCAATGCTTGGCAGCGCTCGCTCGCCA 2739

```

RESULT 21
AB222976/c
ID AB222976 standard; cDNA; 7987 BP.
XX
AC AB222976;
XX
DT 10-APR-2003 (first entry)
XX
DE Eimeria maxima 250 kDa protein encoding cDNA SEQ ID NO:1.
XX
KW Eimeria maxima; 250 kDa antigen; sporozoite; merozoite; infection;
XX vaccine; antiparasitic; gene therapy; immunodominant; gene; ss.
XX
OS Eimeria maxima.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 3..7985
FT FT /*tag= a
FT FT /partial
FT FT /product= "250 kDa antigen protein"
FT FT /note= "the coding sequence is given as shown in the
FT FT figure and any X's in the protein sequence are encoded
FT FT by stop codons; no start or stop codons given"
XX
PN WO2003004684-A2.
XX
XX 16-JAN-2003.
XX
PF 03-JUL-2002; 2002WO-US21237.
XX
XX 06-JUL-2001; 2001US-303670P.
XX
PA (WITC/) WITCOMBE D.
PA (SMIT/) SMITH N C.
PA (WALL/) WALLACH M.
XX
PI Witcombe D, Smith NC, Wallach M;
XX
XX WPI; 2003-201556/19.
XX P-PSDB; ABP56958.
XX
XX New nucleic acid comprising a sequence encoding a 250 kDa polypeptide
XX from sporozoites/merozoites of Eimeria maxima, useful for preparing a
XX vaccine against Eimeria infection
XX
XX Example 3; Fig 6; 198pp; English.
XX
XX The present invention describes a nucleic acid (I) comprising a
XX sequence encoding a 250 kDa polypeptide from Sporozoites/Merozoites of
XX Eimeria maxima or its homologue or complement. Also described: (1) a
XX vector comprising the nucleic acid; (2) a host cell comprising the
XX cell comprising the nucleic acid; (5) producing a recombinant 250 kDa
XX polypeptide; (6) a recombinant polypeptide; (7) a vaccine against
XX E. tenella, E. acervulina, E. necatrix, E. praecox, E. maxima, E. mitis
XX or E. brunetti or a microorganism expressing an immunologically
XX cross-reactive antigen; (8) immunising a subject against infection by
XX Eimeria, or a microorganism expressing an immunologically cross-reactive
XX antigen; (9) conferring upon a newborn subject of an avian species
XX maternal immunity against infection by Eimeria; (10) a fertilised egg
XX from an avian species having an air sac, where the air sac is inoculated
XX with the vaccine; and (11) reducing the output of Eimeria oocytes in
XX faeces from a newborn subject of an avian species. (I) has antiparasitic
XX activity and can be used in a vaccine and in gene therapy. The nucleic
XX acid is useful for preparing a vaccine against E. tenella, E. acervulina,
XX E. necatrix, E. praecox, E. maxima, E. mitis or E. brunetti infection.
XX The present sequence encodes a 250 kDa antigen protein, which is used
XX in an example from the present invention.
XX
XX Sequence 7987 BP; 2022 A; 2217 C; 2370 G; 1378 T; 0 other;

```

```

Query Match 45.8%; Score 23.8; DB 25; Length 7987;
Best Local Similarity 72.1%; Pred. No. 1.3e+02;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 7 TTCTCTCAGCTCTCATCGCTCTCGTGGCTGGCGGCTCGCCA 49
Db 3011 TTCTTCTGGCAGCTCATCAATGTCTTGGCAGCTCGCTCGCCA 2969

RESULT 22
AB222984/c
ID AB222984 standard; cDNA; 7990 BP.
XX
AC AB222984;
XX
DT 10-APR-2003 (first entry)
XX
DE E. maxima immunodominant protein encoding cDNA SEQ ID NO:25.
XX
KW Eimeria maxima; 250 kDa antigen; sporozoite; merozoite; infection;
XX vaccine; antiparasitic; gene therapy; gene; ss.
XX
OS Eimeria maxima.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 231..7313
FT FT /*tag= a
FT FT /product= "immunodominant protein"
FT FT sig_peptide 231..308
FT FT /*tag= b
FT FT mat_peptide 309..7310
FT FT /*tag= c
XX
XX WO2003004684-A2.
XX
XX 16-JAN-2003.
XX
XX 03-JUL-2002; 2002WO-US21237.
XX
XX 06-JUL-2001; 2001US-303670P.
XX
XX (WITC/) WITCOMBE D.
XX (SMIT/) SMITH N C.
XX (WALL/) WALLACH M.
XX
XX Witcombe D, Smith NC, Wallach M;
XX
XX WPI; 2003-201556/19.
XX P-PSDB; ABP56972.
XX
XX New nucleic acid comprising a sequence encoding a 250 kDa polypeptide
XX from sporozoites/merozoites of Eimeria maxima, useful for preparing a
XX vaccine against Eimeria infection
XX
XX Example 5; Fig 29; 198pp; English.
XX
XX The present invention describes a nucleic acid (I) comprising a
XX sequence encoding a 250 kDa polypeptide from Sporozoites/Merozoites of
XX Eimeria maxima or its homologue or complement. Also described: (1) a
XX vector comprising the nucleic acid; (2) a host cell comprising the
XX cell comprising the nucleic acid; (5) producing a recombinant 250 kDa
XX polypeptide; (6) a recombinant polypeptide; (7) a vaccine against
XX E. tenella, E. acervulina, E. necatrix, E. praecox, E. maxima, E. mitis
XX or E. brunetti or a microorganism expressing an immunologically
XX cross-reactive antigen; (8) immunising a subject against infection by
XX Eimeria, or a microorganism expressing an immunologically cross-reactive
XX antigen; (9) conferring upon a newborn subject of an avian species
XX maternal immunity against infection by Eimeria; (10) a fertilised egg
XX from an avian species having an air sac, where the air sac is inoculated
XX with the vaccine; and (11) reducing the output of Eimeria oocytes in
XX faeces from a newborn subject of an avian species. (I) has antiparasitic
XX activity and can be used in a vaccine and in gene therapy. The nucleic
XX acid is useful for preparing a vaccine against E. tenella, E. acervulina,
XX E. necatrix, E. praecox, E. maxima, E. mitis or E. brunetti infection.
XX The present sequence encodes a 250 kDa antigen protein, which is used
XX in an example from the present invention.
XX
XX Sequence 7987 BP; 2022 A; 2217 C; 2370 G; 1378 T; 0 other;

```



```

XX OS Streptomyces verticillus.
XX FH Key
XX FT CDS Location/Qualifiers
XX FT     223..564
XX FT     /tag= a
XX FT     /transl_except= (pos: 1..3, aa: Met)
XX FT     /note= "ORF 30; encodes AAB07556"
XX FT     561..2309
XX FT     /tag= b
XX FT     /transl_except= (pos: 1..3, aa: Met)
XX FT     /note= "ORF 29; encodes AAB07557"
XX FT     2767..3486
XX FT     /tag= c
XX FT     /note= "ORF 28; encodes AAB07558"
XX FT     3527..5593
XX FT     /tag= d
XX FT     /transl_except= (pos: 1..3, aa: Met)
XX FT     /note= "ORF 27; encodes AAB07559"
XX FT     5806..12294
XX FT     /tag= e
XX FT     /note= "ORF 26; encodes AAB07560"
XX FT     12291..15491
XX FT     /tag= f
XX FT     /note= "ORF 25; encodes AAB07561"
XX FT     15488..21013
XX FT     /tag= g
XX FT     /note= "ORF 24; encodes AAB07562"
XX FT     21010..24666
XX FT     /tag= h
XX FT     /transl_except= (pos: 1..3, aa: Met)
XX FT     /note= "ORF 23; encodes AAB07563"
XX FT     24663..32690
XX FT     /tag= i
XX FT     /note= "ORF 22; encodes AAB07564"
XX FT     32893..34830
XX FT     /tag= j
XX FT     /note= "ORF 21; encodes AAB07565"
XX FT     34827..35804
XX FT     /tag= k
XX FT     /transl_except= (pos: 1..3, aa: Met)
XX FT     /note= "ORF 20; encodes AAB07566"
XX FT     35918..37302
XX FT     /tag= l
XX FT     /transl_except= (pos: 1..3, aa: Met)
XX FT     /note= "ORF 19; encodes AAB07567"
XX FT     37299..39215
XX FT     /tag= m
XX FT     /transl_except= (pos: 1..3, aa: Met)
XX FT     /note= "ORF 18; encodes AAB07568"
XX FT     39301..47181
XX FT     /tag= n
XX FT     /note= "ORF 17; encodes AAB07569"
XX FT     47178..49985
XX FT     /tag= o
XX FT     /transl_except= (pos: 1..3, aa: Met)
XX FT     /note= "ORF 16; encodes AAB07570"
XX FT     49982..51001
XX FT     /tag= p
XX FT     /transl_except= (pos: 1..3, aa: Met)
XX FT     /note= "ORF 15; encodes AAB07571"
XX FT     50998..52386
XX FT     /tag= q
XX FT     /transl_except= (pos: 1..3, aa: Met)
XX FT     /note= "ORF 14; encodes AAB07572"
XX FT     52383..52946
XX FT     /tag= r
XX FT     /note= "ORF 13; encodes AAB07573"
XX FT     53018..54190
XX FT     /tag= s
XX FT     /note= "ORF 12; encodes AAB07574"
XX FT     54187..55824
XX FT     /tag= t

```

```

FT CDS /note= "ORF 11; encodes AAB07575"
FT FT 55821..56093
FT FT /*tag= u
FT FT /transl_except= (pos: 1..3, aa: Met)
FT FT /note= "ORF 10; encodes AAB07576"
FT FT 56090..57586
FT FT /*tag= v
FT FT /transl_except= (pos: 1..3, aa: Met)
FT FT /note= "ORF 9; encodes AAB07577"
FT FT 57583..58857
FT FT /*tag= w
FT FT /note= "ORF 8; encodes AAB07578"
XX XX
XX PN W0200040704-A1.
XX XX 13-JUL-2000.
XX PD
XX PF 06-JAN-2000; 2000WO-US00445.
XX XX
XX PR 06-JAN-1999; 99US-0115435.
XX PR 05-FEB-1999; 99US-0118848.
XX PR 05-JAN-2000; 2000US-0477962.
XX XX
XX PA (REGC ) UNIV CALIFORNIA.
XX XX
XX PI Shen B, Du L, Sanchez C, Chen M, Edwards DJ;
XX XX
XX DR WPI; 2000-465974/40.
XX DR P-PSDB; AAB07556, AAB07557, AAB07558, AAB07559, AAB07560, AAB07561,
XX DR AAB07562, AAB07563, AAB07564, AAB07565, AAB07566, AAB07567, AAB07568,
XX DR AAB07569, AAB07570, AAB07571, AAB07572, AAB07573, AAB07574, AAB07575,
XX DR AAB07576, AAB07577, AAB07578.
XX XX
XX PT New bleomycin gene cluster components useful for peptide and/or
XX PT polyketide metabolites, especially bleomycin, production and for
XX PT chemically modifying biological molecules -
XX XX
XX PS Claim 8; Page 97-136; 162pp; English.
XX CC
XX CC The present sequence represents the BLM (Bleomycin) gene cluster,
XX CC containing open reading frames (ORFs) 8-30. The proteins encoded
XX CC by the gene cluster are useful for producing peptides and/or polyketide
XX CC metabolites, especially bleomycin or bleomycin analogues. They are
XX CC also useful for chemically modifying biological molecules to produce
XX CC branched methyl groups, and for coupling amino acids and fatty
XX CC acids. They may be reacted with an apo-carrier protein and coenzyme A
XX CC to produce a holo-carrier protein. The BLM gene cluster or catalytic
XX CC domains can be used individually or collectively to produce
XX CC thiazolidine, thiazoline, bithiazoline and bithiazoline-containing
XX CC microbial metabolites. The BLM gene cluster may also be used to produce
XX CC sugars.
XX SQ Sequence 58857 BP; 7256 A; 25139 C; 19353 G; 7109 T; 0 other;
XX
XX Query Match 45.8%; Score 23.8; DB 21; Length 58857;
XX Best Local Similarity 66.7%; Pred. No. 1.7e+02;
XX Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
XX
XX QY 2 TGCAGTTCTCTCACGCTCTCATGCTCTGCTGCTCGCGCCCTCGCCAGTG 52
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX Db 10310 TGCAGTTCTCTCCGACGCCCTCATGAGAGGCTCCGACACAGCTAGCCCGCG 10360
XX
XX RESULT 25
XX AAH68534
XX ID AAH68534 standard; DNA; 309400 BP.
XX XX
XX AC AAH68534;
XX XX
XX DT 26-SEP-2001 (first entry)
XX XX
XX DE C glutamicum coding sequence fragment SEQ ID NO: 7069.
XX FT

```

KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 XX organic acid synthesis; ds.
 OS Corynebacterium glutamicum.
 XX
 PN EPI108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000EP-0127688.
 XX
 PR 16-DEC-1999; 99JP-0377484.
 PR 07-APR-2000; 2000JP-0159162.
 PR 03-AUG-2000; 2000JP-0280988.
 XX
 PA (KYOWA) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 DR WPI; 2001-376931/40.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 XX
 PS Disclosure; SEQ ID NO: 7069; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 XX
 SQ Sequence 309400 BP; 70133 A; 86477 C; 83115 G; 69675 T; 0 other;
 Query Match 45.8%; Score 23.8; DB 22; Length 309400;
 Best Local Similarity 66.7%; Pred. No. 2e-02;
 Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
 OY 2 TGCAGTTCTCTACGCTCTCATCGCTCTCGCTCGCTCGCTCGCTCGCTCGCTCG 52
 Db 218393 TTCAGATCGGATACGCCATCATCGGTATCGCGCAGCAGGTCTCGCGCTG 218443
 RESULT 26
 ABZ11723
 ID ABZ11723 standard; cDNA; 1415 BP.
 XX
 AC ABZ11723;
 XX
 XX 20-JAN-2003 (first entry)
 DT
 XX
 DE Human polynucleotide SEQ ID NO 605.
 XX
 KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 KW arthritis; cytostatic; immunomodulator; nontropic; neuroprotective;
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 KW haemostatic; vulneratic; fungicide; antibacterial; virucide; protozoacide;
 KW antiarthritic; gene; ss.

OS Homo sapiens.
 XX
 PN WQ200270539-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 05-MAR-2002; 2002WO-US05095.
 XX
 PR 05-MAR-2001; 2001US-0799451.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 XX
 DR WPI; 2002-759812/82.
 DR P-PSDB; ABP69506.
 XX
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative, or
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
 PT platelet or coagulation disorders
 XX
 PS Claim 1: SEQ ID NO 605; 1012pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences
 CC (ABZ1119-ABZ12066) or their mature protein coding portion, active domain
 CC coding protein or complementary sequences. The polynucleotides are useful
 CC for identifying expressed genes or for physical mapping of human genome.
 CC The encoded polypeptides (ABP6902-ABP69849) are useful as molecular
 CC weight markers, as a food supplement, for generating antibodies, in
 CC medical imaging, screening and diagnostic assays and for treating
 CC cell-proliferative disorders (cancer), neurodegenerative diseases
 CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
 CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
 CC disorders, platelet or coagulation disorders, wound, burns, incision,
 CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
 CC parasitic), arthritis, etc.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1415 BP; 337 A; 490 C; 386 G; 202 T; 0 other;
 Query Match 45.4%; Score 23.6; DB 24; Length 1415;
 Best Local Similarity 69.6%; Pred. No. 1.3e+02;
 Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 OY 1 ATGCAGTTCTCTACGCTCTCATCGCTCTCGCTCGCTCGCTCGCTCG 46
 Db 29 ATGCAGTTCTCTCTCGCTCTGTGCGCCCGCTCTCGCTCGCTCGCTCG 74
 RESULT 27
 ABX34289/c
 ID ABX34289 standard; DNA; 135638 BP.
 XX
 AC ABX34289;
 XX
 XX 11-FEB-2003 (first entry)
 DT
 XX
 DE S. atroolivaceus leinamycin biosynthesis gene cluster.
 XX
 KW Leinamycin biosynthesis gene cluster; Lmn; open reading frame; ORF;
 KW anti-tumour antibiotic; broad spectrum antimicrobial activity;
 KW Gram-positive; Gram-negative bacteria; chemical modification;
 KW metabolite; apo-carrier protein; holo-carrier protein; tumour;
 KW polypeptide; hybrid polypeptide/polyketide metabolite; Lmn production;
 KW cytostatic; gene; ds.
 XX
 OS Streptomyces atroolivaceus.

```
XX PN WO200277179-A2.
XX PD 03-OCT-2002.
XX PF 22-MAR-2002; 2002WO-US08937.
XX PR 26-MAR-2001; 2001US-278935P.
XX PA (REGC ) UNIV CALIFORNIA.
XX PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX PI Shen B, Cheng Y, Tang G;
XX DR WPI; 2003-018907/01.
XX DR P-PSDR; ABU11341, ABU11342, ABU11343, ABU11344, ABU11345, ABU11346,
XX DR ABU11347, ABU11348, ABU11349, ABU11350, ABU11351, ABU11352, ABU11353,
XX DR ABU11354, ABU11355, ABU11356, ABU11357, ABU11358, ABU11359, ABU11360,
XX DR ABU11361, ABU11362, ABU11363, ABU11364, ABU11365, ABU11366, ABU11367,
XX DR ABU11368, ABU11369, ABU11370, ABU11371, ABU11372, ABU11373, ABU11374,
XX DR ABU11375, ABU11376, ABU11377, ABU11378, ABU11379, ABU11380, ABU11381,
XX DR ABU11382, ABU11383, ABU11384, ABU11385, ABU11386, ABU11387, ABU11388,
XX DR ABU11389, ABU11390, ABU11391, ABU11392, ABU11393, ABU11394, ABU11395,
XX DR ABU11396, ABU11397, ABU11398, ABU11399, ABU11400, ABU11401, ABU11402,
XX DR ABU11403, ABU11404, ABU11405, ABU11406, ABU11407, ABU11408, ABU11409,
XX DR ABU11410, ABU11411.
XX PT Novel gene cluster responsible for synthesis of leinamycin in
XX PT Streptomycetes atroolivaceus useful for making various peptide and/or
XX PT polypeptide, and/or hybrid polypeptide/polyketide metabolites -
XX PS Claim 6; Page 81-127; 185pp; English.
XX CC The present invention relates to the isolation of the Streptomycetes
XX CC atroolivaceus leinamycin (Lnm) biosynthesis gene cluster containing
XX CC 71 open reading frames (ORFs) -35 through -1, ORFs lnmA through
XX CC lnmZ, and ORFs +1 through +9. Leinamycin is a novel anti-tumour
XX CC antibiotic produced by several Streptomycetes species. It exhibits
XX CC broad spectrum antimicrobial activity against Gram-positive and
XX CC Gram-negative bacteria, but not against fungi. The polypeptides encoded
XX CC by the Lnm biosynthesis gene cluster ORFs are useful for chemically
XX CC modifying a molecule in a host cell. The host cell is a bacterium or
XX CC eukaryotic cell, including a mammalian, yeast, plant, fungal, or insect
XX CC cell. The molecule is an endogenous metabolite produced by the host
XX CC cell or exogenously supplied metabolite, or an amino acid, and the
XX CC polypeptide is a peptide synthetase or amino transferase. The
XX CC polypeptides encoded by the Lnm gene cluster are useful for converting
XX CC an apo-carrier protein to a holo-carrier protein. Lnm shows potent
XX CC antitumour activity in tumour models in vivo. The Lnm gene cluster
XX CC modules and/or catalytic domains are useful for making various peptide
XX CC and/or polypeptide, and/or hybrid polypeptide/polyketide metabolites.
XX CC The proteins encoded by the ORFs are useful alone, or in combination
XX CC with other active domains to modify various target substrates. The
XX CC Lnm gene cluster is useful to upregulate endogenous Lnm production to
XX CC permit Lnm production in cells and/or to make various modified Lnm.
XX CC Lnm, its analogue, or other polypeptide, peptide or hybrid
XX CC polypeptide/peptide metabolites are useful as therapeutic agents, to
XX CC treat a number of disorders, depending upon the type of metabolites.
XX CC The present sequence represents the S. atroolivaceus leinamycin
XX CC biosynthesis gene cluster.
XX SQ Sequence 135638 BP; 18570 A; 49096 C; 49039 G; 18933 T; 0 other;
Query Match 45.4%; Score 23.6; DB 25; Length 135638;
Best Local Similarity 76.3%; Pred. No. 2.1e+02;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 15 CGCTCTCATCTCTCGTCCGCTCGCGCCCTCGCCAGTG 52
|||||
DB 34257 CGCTCTCGCGCGCTCGCGCGCGCCCTCGTCTCGCCGCTG 34220
|||||
RESULT 28
```

```
AAJ38151
ID AAD38151 standard; DNA; 145 BP.
XX AC AAD38151;
XX DT 10-SEP-2002 (first entry)
XX DE pCON3 deletion mutant construct intron A fragment.
XX KW Cytomegalovirus; CMV; expression construct; gene therapy; cytostatic;
XX KW human cytomegalovirus; hCMV; malignant cancer; immunisation; intron A;
XX OS ds.
XX OS Unidentified.
XX PN WO200231137-A2.
XX PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US32050.
XX PR 13-OCT-2000; 2000US-240502P.
XX PA (CHIR ) CHIRON CORP.
XX PI Thudium K, Selby M, Ulmer J;
XX DR WPI; 2002-452346/48.
XX PT Novel human cytomegalovirus Intron A fragment for use in expression
XX PT constructs, lacks full-length Intron A sequence, and enhance expression
XX PT levels when present in expression constructs -
XX PS Disclosure; Page 43; 44pp; English.
XX CC The invention relates to cytomegalovirus (CMV) intron A fragments for
XX CC use in expression construct for expressing gene products. Recombinant
XX CC expression construct is useful for producing a recombinant polypeptide.
XX CC Human cytomegalovirus (hCMV) intron A fragment is useful in expression
XX CC constructs to express a wide variety of substances including peptides.
XX CC It is also useful for producing proteins useful for treating a variety
XX CC of malignant cancers, and for producing proteins useful for prevention,
XX CC treatment and/or diagnosis of a wide variety of diseases. Recombinant
XX CC expression construct is used in nucleic acid immunisation and gene
XX CC therapy. The present sequence is pCON3 deletion mutant construct
XX CC intron A fragment used in the invention.
XX SQ Sequence 145 BP; 27 A; 46 C; 27 G; 45 T; 0 other;
Query Match 45.0%; Score 23.4; DB 24; Length 145;
Best Local Similarity 67.3%; Pred. No. 1.1e+02;
Matches 33; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1 ATGCAGTCTCTCACGCTCTCATCGCTCTCGTCTCGTCCGCGCCCTCGCCA 49
|||||
DB 25 AGGCACACCCCTTTGGCTCTTATGCATCTCGTCTCGCGCGCGCCA 73
|||||
RESULT 29
ABN20845
ID ABN20845 standard; cDNA; 450 BP.
XX AC ABN20845;
XX DT 24-JUN-2002 (first entry)
XX DE Human ORFX polynucleotide sequence SEQ ID NO:10167.
XX KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX KW hypertension; hypothyroidism; cholesterol ester storage disease;
```

KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis; gene; ss.
 XX Homo sapiens.
 PN WO200192523-A2.
 XX 06-DEC-2001.
 PD
 XX 29-MAY-2001; 2001WO-US10836.
 PF
 XX 30-MAY-2000; 2000US-206132P.
 PR 29-AUG-2000; 2000US-228716P.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shinkets RA, Leach MD;
 PI
 XX WPI; 2002-106308/14.
 DR P-PSDB; ABP05093.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders -
 XX
 XX Disclosure; SEQ ID 10167; 1037pp; English.
 PS
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP0010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 450 BP; 86 A; 106 C; 123 G; 133 T; 2 other;
 Query Match 45.0%; Score 23.4; DB 24; Length 450;
 Best Local Similarity 81.8%; Pred. No. 1.3e+02;
 Matches 27; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 15 CGCTCTCATCGCTCTGCTGCTGCGCTGCGGCTCGC 47
 |||||
 Db 145 CGCTCTCATGCGCTGCTGCTGCTGCTGCTGCTGCTG 177
 |||||
 RESULT 30
 AAA94737
 ID AAA94737 standard; DNA; 855 BP.
 XX
 AC AAA94737;
 XX
 DT 02-FEB-2001 (first entry)
 XX

DE Corn auxin transport protein cione cillc.pk001.b7 DNA sequence.
 XX
 KW Auxin transport protein; corn;
 KW root development; gene mapping; plant breeding; herbicide; ss.
 XX
 OS Zea mays.
 XX
 PN WO200068389-A2.
 XX
 PD 16-NOV-2000.
 XX
 PF 03-MAY-2000; 2000WO-US12061.
 XX
 PR 07-MAY-1999; 99US-0133040.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 XX Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;
 PI
 XX WPI; 2000-687647/67.
 DR P-PSDB; AAB26952.
 XX
 XX New nucleic acid sequences encoding new auxin transport proteins,
 PT useful for modulating root growth of plants and to screen for
 PT herbicides -
 XX
 XX Claim 2; Page 93; 94pp; English.
 PS
 XX Auxins are plant hormones that influence plant behaviour and development
 CC e.g. vascular tissue differentiation, apical development, tropic
 CC responses and organ (e.g. flower, leaf) development. The present
 CC invention relates to corn auxin transport protein coding sequences and
 CC proteins. The present sequence is one such coding sequence. This sequence
 CC may be used to modulate root development, e.g. to produce a more robust
 CC root system, alter root angle or redirect root growth. Also, the present
 CC sequence may be useful for gene mapping (e.g. for plant breeding) and to
 CC identify loss of function mutants. The protein encoded by the present
 CC sequence may be useful for raising specific antibodies, for the detection
 CC of auxin transport proteins and to design and/or identify specific
 CC inhibitors of auxin transport proteins, potentially useful as herbicides.
 XX
 SQ Sequence 855 BP; 152 A; 308 C; 266 G; 129 T; 0 other;
 Query Match 45.0%; Score 23.4; DB 21; Length 855;
 Best Local Similarity 73.2%; Pred. No. 1.4e+02;
 Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 12 TCACGCTCTCATCGCTCTGCTGCTGCGCTGCGGCTCGCCAGTG 52
 |||||
 Db 80 TCACGCTCTCTCTCCCTGTCGACGCTGCCAACCACGCTGGTG 120
 |||||
 RESULT 31
 AAC55847
 ID AAC55847 standard; DNA; 2466 BP.
 XX
 AC AAC55847;
 XX
 DT 19-JAN-2001 (first entry)
 XX
 XX Mitomycin biosynthetic gene cluster related alpha amylase orf20.
 DE
 XX Mitomycin; biosynthesis; mitosome ring system; antibiotic; anti-cancer;
 KW anti-inflammatory; immune-enhancer; immunosuppressant; asthma;
 KW chronic obstructive pulmonary disease; respiratory inflammation;
 KW fungicide; pesticide; ds.
 XX
 OS Streptomyces lavendulae.
 XX
 PN WO200053737-A2.
 XX
 PD 14-SEP-2000.

```
XX 10-MAR-2000; 2000WO-US06394.
XX
XX
XX
XX 12-MAR-1999; 99US-0266965.
XX
XX (MINU ) UNIV MINNESOTA.
XX (SHER/) SHERMAN D H.
XX (MAOY/) MAO Y.
XX (VARO/) VAROGLU M.
XX (HEMM/) HE M.
XX (SHEL/) SHELTON P C.
XX
XX Sherman DH, Mao Y, Varoglu M, He M, Sheldon PC;
XX
XX WPI; 2000-601980/57.
XX
XX Novel nucleic acid molecule comprising mitomycin biosynthetic gene
XX cluster useful for cloning mitomycin biosynthetic genes for elucidating
XX the molecular basis of mitosane ring system biosynthesis -
XX
XX Example 1; Page 324-325; 399pp; English.
XX
XX This invention relates to isolated and purified nucleic acid molecules
XX from the mitomycin biosynthetic gene cluster. Mitomycins are a group of
XX natural products that contain a variety of functional groups, including
XX amino benzoquinone and axiridine ring systems. The S. lavendulae
XX mitomycin biosynthetic gene cluster comprises 47 mitomycin genes
XX spanning 55kb of DNA. The invention includes an expression cassette
XX comprising a mitomycin biosynthetic gene operably linked to a promoter,
XX and host cells transformed with the cassette. The nucleotide, and protein
XX sequences and the transformed host cells of the invention result in
XX antitastmatic, antiinflammatory, cytostatic, immunomodulatory, and
XX antibiotic activities. The nucleotide sequences are used to elucidate the
XX molecular basis for the biosynthesis of the mitosane ring system, as well
XX as to engineer the biosynthesis of novel natural products, e.g.
XX antibiotics, anti-inflammatory agents, anti-cancer agents,
XX immune-enhancers, immunosuppressants, agents to treat asthma, chronic
XX obstructive pulmonary disease as well as other disease involving
XX respiratory inflammation, or cholesterol-lowering agents or as crop
XX protection agents (e.g. fungicides or insecticides) as well as
XX biopolymers, e.g., in packaging or biomedical applications, or to engineer
XX PHA monomer syntheses. Sequences AAC55782-C55881, AAC55815-C55849 and
XX AAB32485-B32542 represent mitomycin biosynthetic gene cluster DNA
XX sequences and encoded proteins. Sequences AAC55812-C55814,
XX AAC55850-C55856 and AAC55862-C55869 represent PCR primers used in the
XX cloning of the mitomycin biosynthetic genes.
XX
XX Sequence 2466 BP; 460 A; 860 C; 822 G; 324 T; 0 other;
XX
XX Query Match 45.0%; Score 23.4; DB 21; Length 2466;
XX Best Local Similarity 73.2%; Pred. No. 1.6e+02;
XX Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
XX
XX 9 CTCTCAGCTCTCATCGCTCTCGTCTCGCTCGCGCTCGCGCA 49
XX ||||| ||||| ||||| ||||| ||||| |||||
XX 48 CGCTGCGCGCTCATGCGCTCTCGTCTCGCTCGCGCACCGCA 88
XX
XX RESULT 32
XX ABL13923
XX ID ABL13923 standard; cDNA; 2840 BP.
XX
XX ABL13923;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 36251.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX PI Venter JC, Adams M, He M, Varoglu M, Myers EW;
XX
```

```
PN WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX P-PSDB; ABB69820.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 36251; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (AAB5737-ABB72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 2840 BP; 671 A; 724 C; 723 G; 722 T; 0 other;
XX
XX Query Match 45.0%; Score 23.4; DB 23; Length 2840;
XX Best Local Similarity 73.2%; Pred. No. 1.6e+02;
XX Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
XX
XX 10 TCTCAGCTCTCATCGCTCTCGTCTCGCTCGCGCTCGCGCA 50
XX ||||| ||||| ||||| ||||| ||||| |||||
XX 2774 TCCTCCCGATCTCTGGTGTCTCTCGTCTCGCGCTCGCGCG 2814
XX
XX RESULT 33
XX ABL08443/c
XX ID ABL08443 standard; cDNA; 5736 BP.
XX
XX ABL08443;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 19811.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
```

```
DR WPI; 2001-65860/75.  
DR P-PSDB; ABB64340.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
PS Claim 1; SEQ ID NO 19811; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published_pct_sequences.  
XX  
SQ Sequence 5736 BP; 1503 A; 1718 C; 1588 G; 927 T; 0 other;  
  
Query Match      45.0%; Score 23.4; DB 23; Length 5736;  
Best Local Similarity    67.3%; Pred.No. 1.7e+02;  
Matches   33; Conservative     0; Mismatches 16; Indels   0; Gaps  
  
OY      4 CAGTCTCCTCACGGTCCTCATCGCTCTGCTGCCTGCCGCCATGCACGTG S2  
        |||||          |||||         |||||       |||||      |||||  
Db      5544 CAGTTCTTGCAAGCTCTCTGTCCTCTCTCTCTCTCTGTCATCGTCGTG 5496  
  
RESULT 34  
ABLI3922  
ID      ABLI3922 standard; cDNA; 6782 BP.  
XX  
AC      ABLI3922;  
XX  
DT      26-MAR-2002 (first entry)  
XX  
DE      Drosophila melanogaster expressed polynucleotide SEQ ID NO 36248.  
XX  
KW      Drosophila; developmental biology; cell signalling; insecticide;  
KM      pharmaceutical; gene; ss.  
XX  
OS      Drosophila melanogaster.  
XX  
PN      WO200171042-A2.  
XX  
PD      27-SEP-2001.  
XX  
PF      23-MAR-2001; 2001WO-US09231.  
XX  
PR      23-MAR-2000; 2000US-191637P.  
PR      11-JUL-2000; 2000US-0614150.  
XX  
PA      (PEKE ) PE CORP NY.  
XX  
PI      Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR      WPI; 2001-65860/75.  
DR      P-PSDB; ABB69819.  
XX  
PT      New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Claim 1; SEQ ID NO 36248; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of
```



```
Query Match      44.6%; Score 23.2; DB 25; Length 271;
Best Local Similarity 70.5%; Pred. No. 1.4e+02;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY      8 TCTCTCAGCTCTCATCGCTCTCGTCTCGCTGCGCGGCTCGCCAGT 51
DB      59 TCTGTCTGTCTCTCATCGCTCTCATCTCCATTGTGTGGCCAGCCAGT 16

RESULT 38
ABQ98305/c
ID ABQ98305 standard; DNA; 450 BP.
XX
AC ABQ98305;
XX
DT 04-NOV-2002 (first entry)
XX
DE Human ORF112 coding sequence.
XX
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN US2002082206-A1.
XX
PD 27-JUN-2002.
XX
PF 30-MAY-2001; 2001US-0867550.
XX
PR 30-MAY-2000; 2000US-208427P.
XX
PA (LEAC/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA (CONL/) CONLEY P B.
PA (TOPP/) TOPPER J N.
PA (LAWD/) LAW D.
XX
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
WPI; 2002-626554/67.
DR P-PSDB; ABP63742.
XX
New polypeptide designated ORFX are present in human atherogenic cells
and are useful to prevent and treat ORFX-associated disorders including
cancer, allergy, wound healing or autoimmune, cardiovascular or
inflammatory disease -
XX
Claim 2; SEQ ID 223; 78pp; English.
XX
The present invention relates to novel human ORFX polypeptides and their
coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
were discovered in human atherogenic cells, in particular in platelets
and human umbilical vein endothelial cells (HUVEC) and are expressed in
many other tissues as well. Atherogenic cells are cells which have the
potential to develop atherosclerotic plaques. The ORFX polypeptides and
nucleic acids are useful for treating or preventing a pathological
condition associated with an ORFX-associated disorder, e.g. cancer,
cardiovascular disease, allergy, autoimmune disease, wound healing, blood
coagulation disorders or inflammatory disorders.
XX
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from the
USFTO web site at seqdata.uspto.gov/sequence.html?DocID=2002082206.
XX
Sequence 450 BP; 113 A; 113 C; 159 G; 62 T; 3 other;

Query Match      44.6%; Score 23.2; DB 24; Length 450;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      13 CACGCTCTCATCGCTCTCGTCTCGCTGCGCGCTCGCC 48
DB      207 CTCGATCTCGCGGCTCTCTGTCGCGCCCTTCTCTCGCC 172

RESULT 39
ABL26959
ID ABL26959 standard; DNA; 867 BP.
XX
AC ABL26959;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 32350.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI; 2001-656860/75.
XX
New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -
XX
Claim 1; SEQ ID NO 32350; 21pp + Sequence Listing; English.
XX
The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
sequences (ABL01840-ABL16175) and the encoded proteins
(ABB57737-ABB72072).
XX
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
XX
Sequence 867 BP; 220 A; 251 C; 207 G; 189 T; 0 other;

Query Match      44.6%; Score 23.2; DB 23; Length 867;
Best Local Similarity 65.4%; Pred. No. 1.6e+02;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY      1 ATGCAGTTTCTCTACGCTCTCATCGCTCTCGCTGCGCGCTCGCCAGTG 52
DB      636 AGGCAGTTTCTCTCGCGGCTCTCTCTCTCTCATGACACCTCGGCATTG 687

RESULT 40
ABL60268
ID ABL60268 standard; cDNA; 1563 BP.
XX
AC ABL60268;
XX
DT 23-AUG-2002 (first entry)
XX
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2003, 07:05:30 ; Search time 57 Seconds
(without alignments)
402.665 Million cell updates/sec

Title: US-10-081-935-1
 Perfect score: 52
 Sequence: 1 atgagttctctcacgctct.....cgtgcgcgcgcctcgccagtq 52

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs. 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : Issued Patents_NA:**
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:**
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:**
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:**
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:**
5: /cgn2_6/ptodata/2/ina/PTCTUS_COMB.seq:**
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:**
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	\$				
1	25	48.1		651	4	US-09-252-991A-10364	Sequence 10364, A
2	25	48.1		2004	4	US-09-252-991A-10609	Sequence 10609, A
C 3	25	48.1		2691	4	US-09-252-991A-10245	Sequence 10245, A
4	23.8	45.8		49377	1	US-08-764-233A-1	Sequence 1, Appli
5	23.4	45.0		972	4	US-09-252-991A-15702	Sequence 15702, A
6	23.4	45.0		975	4	US-09-252-991A-15674	Sequence 15674, A
C 7	23.4	45.0		1885	4	US-09-252-991A-15656	Sequence 15656, A
8	23.4	45.0		2466	4	US-09-266-965-83	Sequence 83, Appl
9	23.4	45.0		18034	4	US-09-266-965-75	Sequence 75, Appl
C 10	23.2	44.6		201	4	US-09-252-991A-782	Sequence 782, App
11	23.2	44.6		1038	1	US-07-661-610C-13	Sequence 13, Appl
12	23.2	44.6		1899	1	US-08-443-104-6	Sequence 6, Appli
13	23.2	44.6		1899	1	US-08-442-859-6	Sequence 6, Appli
14	23.2	44.6		1899	2	US-08-398-489-6	Sequence 6, Appli
15	23.2	44.6		1899	5	PCT-US95-05534-6	Sequence 6, Appli
16	23.2	44.6		2052	1	US-08-443-104-5	Sequence 5, Appli
17	23.2	44.6		2052	1	US-08-238-130-6	Sequence 5, Appli
18	23.2	44.6		2052	1	US-08-442-859-5	Sequence 5, Appli
19	23.2	44.6		2052	2	US-08-398-489-5	Sequence 5, Appli
20	23.2	44.6		2052	2	US-08-894-772-1	Sequence 1, Appli
21	23.2	44.6		2052	2	US-09-207-844-1	Sequence 1, Appli
22	23.2	44.6		2052	4	US-09-252-509-1	Sequence 1, Appli
23	23.2	44.6		2052	5	PCT-US95-05534-5	Sequence 5, Appli
24	23.2	44.6		5437	1	US-07-661-610C-1	Sequence 1, Appli
25	23.2	44.6		5762	4	US-09-647-492A-2	Sequence 2, Appli
C 26	23	44.2		3903	6	5202236-12	Patent No. 5202236
C 27	23	44.2		319608	4	US-09-530-333D-1	Sequence 1, Appli

C	28	44.2	319608	4	US-09-679-409-1	Sequence 1, Appli
	29	22.4	43.1	48	US-08-169-948B-35	Sequence 35, Appli
	30	22.4	43.1	48	US-08-448-873-35	Sequence 35, Appli
	31	22.4	43.1	48	US-08-382-452D-35	Sequence 35, Appli
	32	22.4	43.1	49	US-08-507-362A-19	Sequence 19, Appli
	33	22.4	43.1	702	US-09-146-770-2	Sequence 2, Appli
	34	22.4	43.1	702	US-09-216-295-2	Sequence 2, Appli
	35	22.4	43.1	702	US-09-633-084-2	Sequence 2, Appli
	36	22.4	43.1	702	US-09-075-873-2	Sequence 2, Appli
	37	22.4	43.1	702	US-10-261-997-2	Sequence 2, Appli
	38	22.4	43.1	1050	US-08-032-848C-11	Sequence 11, Appli
	39	22.4	43.1	1050	US-08-438-870-11	Sequence 11, Appli
	40	22.4	43.1	1050	US-08-774-065-1	Sequence 1, Appli
	41	22.4	43.1	1365	US-09-252-991A-8265	Sequence 8265, Ap
	42	22.2	43.7	703	US-08-998-416-751	Sequence 751, App
	43	22.2	43.7	1140	US-08-289-653-2	Sequence 2, Appli
	44	22.2	43.7	2128	US-09-620-312D-197	Sequence 197, App
	45	22.2	43.7	2845	US-08-289-653-1	Sequence 1, Appli

ALIGNMENTS

```

RESULT 1
US-09-252-991A-10364
; Sequence 10364, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10364
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10364

```

```

Query Match      48.1%; Score 25; DB 4; Length 651;
Best Local Similarity 69.4%; Pred. No. 7.1;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      4  CAGTTCCTCAGGCTCTCATGCTCTCGTTCGTCGGGCTCGCAGTG 52
DB      300 CAGCTCGTGTACCTCTCTTGGCTCTCTCGAGCGGAGCGTCAGCG 348

```

```

RESULT 2
US-09-252-991A-10609/c
; Sequence 10609, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10609
; LENGTH: 2004
; TYPE: DNA

```

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10609

Query Match 48.1%; Score 25; DB 4; Length 2004;
Best Local Similarity 69.4%; Pred. No. 8.1;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

4 CAGTTCCTCAGCGCTCATCGCTGCTGCTGCCGCGCTCGCCAGTGG 52
Db 591 CAGCTCGCTGACCTCCTCTTGGCCTGCTGCGACGCGCGCAGCGTACGCG 543

RESULT 3
US-09-252-991A-10245
Sequence 10245, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 10245
LENGTH: 2691
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10245

Query Match 48.1%; Score 25; DB 4; Length 2691;
Best Local Similarity 69.4%; Pred. No. 8.4;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

4 CAGTTCCTCAGCGCTCATCGCTGCTGCTGCCGCGCTCGCCAGTGG 52
Db 2161 CAGCTCGCTGACCTCCTCTTGGCCTGCTGCGACGCGCGCAGCGTACGCG 2209

RESULT 4
US-08-764-233A-1
Sequence 1, Application US/08764233A
Patent No. 5716849
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Schupp, Thomas
APPLICANT: Beck, James J.
APPLICANT: Hill, Dwight S.
APPLICANT: Neff, Snezana
APPLICANT: Ryals, John A.
TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764.233A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/729,214
FILING DATE: 09-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: 1506/CIP6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4937 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Sorangium cellulosum
IMMEDIATE SOURCE:
CLONE: p98/1, pUL3, and pVKM15
FEATURE:
NAME/KEY: misc_feature
LOCATION: 927..19874
OTHER INFORMATION: /product= "Sora"
OTHER INFORMATION: /note= "This gene encodes a protein that is highly homologous to type I PKSs such as eryA from Saccharopolyspora erythraea."
OTHER INFORMATION: are known to be involved in the synthesis of polyketide
OTHER INFORMATION: compounds."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 942..7115
OTHER INFORMATION: /product= "Module 1 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7203..12884
OTHER INFORMATION: /product= "Module 2 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 13455..19616
OTHER INFORMATION: /product= "Module 3 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19870..24556
OTHER INFORMATION: /product= "Module 1 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 24638..30820
OTHER INFORMATION: /product= "Module 2 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 30881..35446
OTHER INFORMATION: /product= "Module 3 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 35528..40114
OTHER INFORMATION: /product= "Module 4 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 40190..46318

```

; OTHER INFORMATION: /product= "Module 5 of SorB"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 46851..47891
; OTHER INFORMATION: /product= "SorM"
; OTHER INFORMATION: /note= "The protein encoded by the sorM gene is highly
; OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
; OTHER INFORMATION: hyposcopus that is involved in the synthesis of the
; OTHER INFORMATION: polyketide rappamycin."
US-08-764-233A-1

Query Match          45.8%; Score 23.4; DB 1; Length 49377;
Best Local Similarity 72.1%; Pred. No. 29;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 6 GTTCTCAGCGCTCATGCTCTCGTGGCTGGCGGCTGCC 48
Db 5048 GTTCACCGCCGCGCTCGCGCTCTCCAGCTGAGCGGCTCACC 5090

RESULT 5
US-09-252-991A-15702
; Sequence 15702; Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15702
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; NAME/KEY: unsure
; LOCATION: (669)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-15702

Query Match          45.0%; Score 23.4; DB 4; Length 972;
Best Local Similarity 73.2%; Pred. No. 25;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 ATGCACTTCTCTCAGCGCTCATGCTCTCGTGGCTGCCG 41
Db 151 ATGCAGGCGCGCGCCGCTGATCGCGCTGTCGGCGCCG 191

RESULT 6
US-09-252-991A-15674
; Sequence 15674; Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15674
; LENGTH: 975

; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; NAME/KEY: unsure
; LOCATION: (571)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-15674

Query Match          45.0%; Score 23.4; DB 4; Length 975;
Best Local Similarity 73.2%; Pred. No. 25;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 ATGCACTTCTCTCAGCGCTCATGCTCTCGTGGCTGCCG 41
Db 53 ATGCAGGCGCGCGCCGCTGATCGCGCTGTCGGCGCCG 93

RESULT 7
US-09-252-991A-15656/c
; Sequence 15656; Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15656
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; NAME/KEY: unsure
; LOCATION: (415)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-15656

Query Match          45.0%; Score 23.4; DB 4; Length 1485;
Best Local Similarity 73.2%; Pred. No. 26;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 ATGCACTTCTCTCAGCGCTCATGCTCTCGTGGCTGCCG 41
Db 933 ATGCAGGCGCGCGCCGCTGATCGCGCTGTCGGCGCCG 893

RESULT 8
US-09-266-965-83
; Sequence 83; Application US/09266965
; Patent No. 6495348
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600.456US1
; CURRENT APPLICATION NUMBER: US/09/266.965
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 08/624,447
; PRIOR FILING DATE: 1996-08-19
; PRIOR APPLICATION NUMBER: PCT/US94/11279
; PRIOR FILING DATE: 1994-10-06
; PRIOR APPLICATION NUMBER: US 08/133,963
; PRIOR FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
```

```

: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 83
: LENGTH: 2466
: TYPE: DNA
: ORGANISM: Streptomyces lavendulae
US-09-266-965-83

```

Query Match	45.0%	Score 23.4	DB 4	length 2466
Best Local Similarity	73.2%	Pred. No. 28		
Matches 30	Conservative 0	Mismatches 11	Indels 0	Gaps 0

Oy 9 CTCTACGCTTCATCGCTCTGTCGGTCCGGCCTGCCA 49
| | | | | | | | | | | | | |
Db 48 CGCTGCGGCCCTGATGGCCCTGTGCGCTCGGCACCGCCA 88

```

US-09-266-965-75
Result 9
US-09-266-965-75
Sequence 75, Application US/09266965
Patent No. 6495348
GENERAL INFORMATION:
APPLICANT: Sherman, D
APPLICANT: Mao, Y
APPLICANT: Varoglu, M
APPLICANT: He, M
APPLICANT: Sheldon, P
TITLE OF INVENTION: Micomycin biosynthetic gene cluster
FILE REFERENCE: 600.4456US1
CURRENT APPLICATION NUMBER: US/09/266,965
CURRENT FILING DATE: 1999-03-12
EARLIER APPLICATION NUMBER: US 08/624,447
EARLIER FILING DATE: 1996-08-19
EARLIER APPLICATION NUMBER: PCT/US94/11279
EARLIER FILING DATE: 1994-10-06
EARLIER APPLICATION NUMBER: US 08/133,963
EARLIER FILING DATE: 1993-10-07
NUMBER OF SEQ ID NOS: 145
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 75
LENGTH: 18034
TYPE: DNA
ORGANISM: Streptomyces lavendulae
FEATURE:
NAME/KEY: unsure
LOCATION: (302)...(302)
OTHER INFORMATION: n is a o r t o r g o c
US-09-266-965-75

```

Query Match	45.0%	Score 23.4	DB 4	Length 18034
Best Local Similarity	73.2%	Pred. No. 35		
Matches	30	Conservative	0	Mismatches 11; Indels 0
Oy	9	CTCTCAGCGTCTTCATCTGCTCTGTCGCTGCGGCGGCTTGCCA	49	
db	11562	CGTTCGCGCGCCGATGAGCCCTCTGTCGCTTCGCGACCGCA	11602	

```

RESULT 10
US-09-252-991A-782/c
; Sequence 782, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

```

```

; SEQ ID NO 782
;
; LENGTH: 201
;
; TYPE: DNA
;
; ORGANISM: Pseudomonas aeruginosa
;
US-09-252-991A-782

```

Query Match	44.6%	Score	23.2	DB	4	Length	201
Best Local Similarity	70.5%	Pred.	No. 24				
Matches	31	Mismatches	13	Indels	0	Gaps	0

Oy 4 CAGTTCCTCAGCGCTCATGCGCTGTGGTGCCGGCGCTGC 47
| | | | | | | | | | | | | | |
Db 135 CTGTTGGCTCATGCTCTGATCGTCTTCGACATGCCCCGTCTGC 92

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

Query Match 44.6%; Score 23.2; DB 1; Length 1030;
Best Local Similarity 65.4%; Pred. No. 29;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Oy 1 ATGAGTTCTTCACGCTTCATCGCTCGTCGCGCGGCTCGCCAGTG 52
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 ACGCGGGTCTCGAAGTCATCATCGGATCGTGCGCCCGGCGTCACAGG 64

RESULT 12
US-08-443-104-6
; Sequence 6, Application US/08443104
; Patent No. 5691162
; GENERAL INFORMATION:

```

APPLICANT: Shuster, Jeffrey R.
APPLICANT: Madden, Mark
APPLICANT: Moyer, Donna L.
APPLICANT: Fuglsang, Claus
APPLICANT: Branner, Sven
TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 56911620 No. 5691162disk of No. 5691162th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Tape
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/443,104
FILING DATE: 17-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/398,489
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 4180.010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1899 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-443-104-6

Query Match 44.6%; Score 23.2; DB 1; Length 1899;
Best Local Similarity 65.4%; Pred. No. 32;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OR 1 ATGCGTTCTCTCAGCCTCTCATGCTCTCGTGGCTGCCGCGCTGCCAGTG 52
DB 1 ATGCGTTCTCTCAGCCTCTCTCTCTCATGCGCTATCCAGCTCGCTGGTG 52

RESULT 13
US-08-442-859-6
Sequence 6, Application US/08442859
Patent No. 5807729
GENERAL INFORMATION:
APPLICANT: Shuster, Jeffrey R.
APPLICANT: Madden, Mark
APPLICANT: Moyer, Donna L.
APPLICANT: Fuglsang, Claus
APPLICANT: Branner, Sven
TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 58077290 No. 5807729disk of No. 5807729th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:

MEDIUM TYPE: Tape
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,859
FILING DATE: 17-MAY-1995
CLASSIFICATION: 252
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/398,489
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 4180.010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1899 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-442-859-6

Query Match 44.6%; Score 23.2; DB 1; Length 1899;
Best Local Similarity 65.4%; Pred. No. 32;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OR 1 ATGCGTTCTCTCAGCCTCTCATGCTCTCGTGGCTGCCGCGCTGCCAGTG 52
DB 1 ATGCGTTCTCTCAGCCTCTCTCTCTCATGCGCTATCCAGCTCGCTGGTG 52

RESULT 14
US-08-398-489-6
Sequence 6, Application US/08398489
Patent No. 5843753
GENERAL INFORMATION:
APPLICANT: Shuster, Jeffrey R.
APPLICANT: Madden, Mark
APPLICANT: Moyer, Donna L.
APPLICANT: Fuglsang, Claus
APPLICANT: Branner, Sven
TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 58437530 No. 5843753disk of No. 5843753th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Tape
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/398,489
FILING DATE: 03-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 4180.010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 1899 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-398-489-6

Query Match 44.6%; Score 23.2; DB 2; Length 1899;
Best Local Similarity 65.4%; Pred. No. 32;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 ATGCACTTCTCAGCCTCTCATGCTCTGTCGCTCGCGGCTCGCCAGTG 52
DB 1 ATGCGTTCTCGAGCTCTCTCTCTCATCGGCTATCCAGCCTCGCTGTG 52

RESULT 15

PCT-US95-05534-6
Sequence 6, Application PC/TUS9505534
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Tape

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05534

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/238/108

FILING DATE: 04-MAY-1994

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/398,489

FILING DATE: 03-MARCH-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.

REGISTRATION NUMBER: 34,086

REFERENCE/DOCKET NUMBER: 4180.204-MO

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 1899 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

PCT-US95-05534-6

Query Match 44.6%; Score 23.2; DB 5; Length 1899;
Best Local Similarity 65.4%; Pred. No. 32;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 ATGCACTTCTCAGCCTCTCATGCTCTGTCGCTCGCGGCTCGCCAGTG 52
DB 1 ATGCGTTCTCGAGCTCTCTCTCTCATCGGCTATCCAGCCTCGCTGTG 52

RESULT 16
US-08-443-104-5
Sequence 5, Application US/08443104
Patent No. 5691162

GENERAL INFORMATION:
APPLICANT: Shuster, Jeffrey R.

APPLICANT: Madden, Mark

APPLICANT: Moyer, Donna L.

APPLICANT: Fuglsang, Claus

APPLICANT: Branner, Sven

TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 56911620 No. 5691162disk of No. 5691162th America, Inc.

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Tape

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/443,104

FILING DATE: 17-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/398,489

FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Agis Dr., Cheryl H.

REGISTRATION NUMBER: 34,086

REFERENCE/DOCKET NUMBER: 4180.010-US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 2052 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-443-104-5

Query Match 44.6%; Score 23.2; DB 1; Length 2052;
Best Local Similarity 65.4%; Pred. No. 32;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 ATGCACTTCTCAGCCTCTCATGCTCTGTCGCTCGCGGCTCGCCAGTG 52
DB 1 ATGCGTTCTCGAGCTCTCTCTCTCATCGGCTATCCAGCCTCGCTGTG 52

RESULT 17

US-08-238-130-6
Sequence 6, Application US/08238130
Patent No. 5702934

GENERAL INFORMATION:

APPLICANT: Hastrup, Sven

APPLICANT: Branner, Sven

APPLICANT: Jorgensen, Birthe R.

APPLICANT: Christensen, Tove

APPLICANT: Jorgensen, Birgitte B.

APPLICANT: Shuster, Jeffrey R.

APPLICANT: Madden, Mark

APPLICANT: Moyer, Donna L.

APPLICANT: Fuglsang, Claus

TITLE OF INVENTION: PROCESSES FOR PRODUCING AN ENZYME
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: No. 57029340 No. 5702934disk of No. 5702934th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Tape
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/238,130
FILING DATE: 04-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 522/93
FILING DATE: 05-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr. Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 3965.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2052 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1..363, 416..802, 856..1821, 1870..2052)
FEATURE:
NAME/KEY: mat peptide
LOCATION: join(1..363, 416..802, 856..1821, 1870..2049)
US-08-238-130-6
Query Match 44.6%; Score 23.2; DB 1; Length 2052;
Best Local Similarity 65.4%; Pred. No. 32;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 ATGCAGTTCTCTCAGCCTCTCATCGCTCTCGTCCGCCGCTCCGCACTG 52
DB 1 ATGCCTTCTCCGACTCTCTCCTCTCATCGGCTATCCAGCTCGCTGTG 52
RESULT 18
US-08-442-859-5
Sequence 5, Application US/08442859
GENERAL INFORMATION:
APPLICANT: Shuster, Jeffrey R.
APPLICANT: Madden, Mark
APPLICANT: Moyer, Donna L.
APPLICANT: Fuglsang, Claus
APPLICANT: Branner, Sven
TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED
TITLE OF INVENTION: ACTIVITY
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: No. 58077290 No. 5807729disk of No. 5807729th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:

MEDIUM TYPE: Tape
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,859
FILING DATE: 17-MAY-1995
CLASSIFICATION: 252
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/398,489
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 4180.010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2052 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-442-859-5
Query Match 44.6%; Score 23.2; DB 1; Length 2052;
Best Local Similarity 65.4%; Pred. No. 32;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 ATGCAGTTCTCTCAGCCTCTCATCGCTCTCGTCCGCCGCTCCGCACTG 52
DB 1 ATGCCTTCTCCGACTCTCTCCTCTCATCGGCTATCCAGCTCGCTGTG 52
RESULT 19
US-08-398-489-5
Sequence 5, Application US/08398489
GENERAL INFORMATION:
APPLICANT: Shuster, Jeffrey R.
APPLICANT: Madden, Mark
APPLICANT: Moyer, Donna L.
APPLICANT: Fuglsang, Claus
APPLICANT: Branner, Sven
TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED
TITLE OF INVENTION: ACTIVITY
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: No. 58437530 No. 5843753disk of No. 5843753th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Tape
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/398,489
FILING DATE: 03-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 4180.010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 2052 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-398-489-5

Query Match 44.6%; Score 23.2; DB 2; Length 2052;
Best Local Similarity 65.4%; Pred. No. 32;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Oy 1 ATGCAGTTCTCTCAGCTTCATGCTCTCGTCCGCGCCCTCGCCAGTG 52
Db 1 ATGCAGTTCTCTCAGCTTCATGCTCTCGTCCGCGCCCTATCCAGCTCGCTGGTG 52

RESULT 20

US-08-894-772-1
Sequence 1, Application US/08894772
Patent No. 5861280

GENERAL INFORMATION:

APPLICANT: Lehmebeck, Jan

TITLE OF INVENTION: Host Cell Expressing Reduced Levels

TITLE OF INVENTION: Of A Metalloprotease And Methods Using The Host Cell In Protei

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 58612800 No. 5861280disk of No. 5861280th America, Inc.

STREET: 405 Lexington Avenue, Suite 6400

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,772

FILING DATE: 27-AUG-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valetta A.

REGISTRATION NUMBER: 35,127

REFERENCE/DOCKET NUMBER: 4300, 204-US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123

TELEFAX: 212-867-0298

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2052 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-894-772-1

Query Match 44.6%; Score 23.2; DB 2; Length 2052;
Best Local Similarity 65.4%; Pred. No. 32;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Oy 1 ATGCAGTTCTCTCAGCTTCATGCTCTCGTCCGCGCCCTCGCCAGTG 52
Db 1 ATGCAGTTCTCTCAGCTTCATGCTCTCGTCCGCGCCCTATCCAGCTCGCTGGTG 52

RESULT 21
US-09-207-844-1

Sequence 1, Application US/09207844
Patent No. 5968774

GENERAL INFORMATION:

APPLICANT: Lehmebeck, Jan

TITLE OF INVENTION: Host Cell Expressing Reduced Levels

TITLE OF INVENTION: Of A Metalloprotease And Methods Using The Host Cell In Protei

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59687740 No. 5968774disk of No. 5968774th America, Inc.

STREET: 405 Lexington Avenue, Suite 6400

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/207,844

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/894,772

FILING DATE: 27-AUG-1997

ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valetta A.

REGISTRATION NUMBER: 35,127

REFERENCE/DOCKET NUMBER: 4300, 204-US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123

TELEFAX: 212-867-0298

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2052 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-09-207-844-1

Query Match 44.6%; Score 23.2; DB 2; Length 2052;
Best Local Similarity 65.4%; Pred. No. 32;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Oy 1 ATGCAGTTCTCTCAGCTTCATGCTCTCGTCCGCGCCCTCGCCAGTG 52
Db 1 ATGCAGTTCTCTCAGCTTCATGCTCTCGTCCGCGCCCTATCCAGCTCGCTGGTG 52

RESULT 22

US-09-252-509-1
Sequence 1, Application US/09252509
Patent No. 6352841

GENERAL INFORMATION:

APPLICANT: Lehmebeck, Jan

TITLE OF INVENTION: No. 6352841el Host Cells and Methods of

TITLE OF INVENTION: Producing Proteins

FILE REFERENCE: 4920,204-US

CURRENT APPLICATION NUMBER: US/09/252,509

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: 1024/96

PRIOR FILING DATE: 1996-09-19

PRIOR APPLICATION NUMBER: PCT/DK97/00397

NUMBER OF SEQ ID NOS: 6

SOFTWARE: FaestSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 2052

TYPE: DNA

ORGANISM: Fusarium oxysporum

FEATURE:
NAME/KEY: mat peptide
LOCATION: (785)...(2049)
NAME/KEY: sig_peptide
LOCATION: (55)...(784)
NAME/KEY: intcon
LOCATION: (364)...(415)
NAME/KEY: intcon
LOCATION: (802)...(854)
NAME/KEY: intcon
LOCATION: (1821)...(1868)
US-09-252-509-1

Query Match 44.6%; Score 23.2; DB 4; Length 2052;
Best Local Similarity 65.4%; Pred. No. 32;
Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 ATGCAGTTCTTCACGCTCTCATCGCTCTGTCGTCGCGGCTCCCACTG 52
Db 1 ATGCGTTCTCCGACTCTCTCCCTCCATCGGCTATCCAGCTCCCTGCTG 52

RESULT 23

PCT-US95-05534-5
Sequence 5, Application PC/TUS9505534

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Novo Nordisk of North America, Inc.

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Tape

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/05534

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/238/108

FILING DATE: 04-MAY-1994

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/398,489

FILING DATE: 03-MARCH-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Agis Dr. Cheryl H.

REGISTRATION NUMBER: 34,086

REFERENCE/DOCKET NUMBER: 4180.204-WO

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2052 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

PCT-US95-05534-5

Query Match 44.6%; Score 23.2; DB 5; Length 2052;
Best Local Similarity 65.4%; Pred. No. 32;

Matches 34; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 ATGCAGTTCTTCACGCTCTCATCGCTCTGTCGTCGCGGCTCCCACTG 52
Db 1 ATGCGTTCTCCGACTCTCTCCCTCCATCGGCTATCCAGCTCCCTGCTG 52

RESULT 24

US-07-661-610C-1

Sequence 1, Application US/07661610C

Patent No. 5292643

GENERAL INFORMATION:

APPLICANT: Shibano, Yuji

APPLICANT: Toyoda, Hideyoshi

APPLICANT: Utsunomiya, Ryutaro

APPLICANT: Obata, Kazuaki

TITLE OF INVENTION: Fusaric Acid Resistant Genes

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,

ADDRESS: P.C.

STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington

STATE: Virginia

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/661,610C

FILING DATE: 19910228

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5292643man F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 2292-010-0

TELEPHONE: (703)521-4500

TELEFAX: (703)486-2347

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5437 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

ORGANISM: Pseudomonas cepacia

FEATURE:

NAME/KEY: RBS

LOCATION: 519..522

FEATURE:

NAME/KEY: CDS

LOCATION: 531..1832

FEATURE:

NAME/KEY: RBS

LOCATION: 1836..1839

FEATURE:

NAME/KEY: CDS

LOCATION: 1847..2275

FEATURE:

NAME/KEY: RBS

LOCATION: 2301..2304

FEATURE:

NAME/KEY: CDS

LOCATION: 2314..3354

FEATURE:

NAME/KEY: RBS

LOCATION: 3405..3409

FEATURE:

```

?      NAME/KEY:  CDS
?      LOCATION:  3414..4040
?
?      FEATURE:
?      NAME/KEY:  RBS
?      LOCATION:  4457..4462
?
?      FEATURE:
?      NAME/KEY:  CDS
?      LOCATION:  4473..4979
?
US-07-661-610C-1

```

Query Match	44.6%	Score 23.2;	DB 1;	Length 5437;
Best Local Similarity	65.4%	Pred. No. 36;		
Matches 34; Conservative	0;	Mismatches 18;	Indels 0;	Gaps 0;

QY 1 ATGAGTTCCTCAGCCTCATGCTCTCGTCGCGCGCGCTCGGCAGTG 52
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2326 ACGCGGCTCTCGAAGTCATCATCGGATCGTGTGGCGCGGCTGTCAGCG 2377

```

RESULT 25
US-09-647-492A-2
; Sequence 2, Application US/09647492A
; Patent No. 6551797
; GENERAL INFORMATION:
; APPLICANT: PFALLER, Rupert
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR PRODUCTION OF PROTEINS
; FILE REFERENCE: PFALLER ET AL.-2 (PC)
; CURRENT APPLICATION NUMBER: US/09/647,492A
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 5762
; TYPE: DNA
; ORGANISM: Trameetes versicolor
US-09-647-492A-2

```

Query Match	44.6%	Score 23.2;	DB 4;	Length 5762;
Best Local Similarity	65.4%;	Pred. No. 36;		
Matches 34;	Conservative 0;	Mismatches 18;	Indels 0;	Gaps 0;

```

Oy      1 ATGCAGTTCCTCAGCGTCATCGCTCTGTCGCTGCGGCGCTGCGCAGTG 52
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      548 ATGTCGAGCTTCACCTCTCTCGCTTTCGTGCTGCTTCCCTTGGCGCTG 592

```

RESULT 266
5202236-12/c
Patent No. 5202236
APPLICANT: MARSH, KATHY J.; ANDERSON, DAVID M.; STRAUSBERG,
SUSAN L.; MCCANDLISH, RUSS; WEI, TENA; FILIPULA, DAVID
TITLE OF INVENTION: METHOD OF PRODUCING BIOADHESIVE
PROTEIN
NUMBER OF SEQUENCES: 39
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/528,762
FILING DATE: 25-MAY-1990
APPLICATION NUMBER: 82,456
FILING DATE: 07-AUG-1987
APPLICATION NUMBER: 933,945
FILING DATE: 24-NOV-1986
APPLICATION NUMBER: 650,128
FILING DATE: 13-SEP-1984
SEQ ID NO.:12
LENGTH: 3903
5202236-12

	Query Match	44.2%	Score 23;	DB 6;	Length 3903;
	Best Local Similarity	74.4%;	Pred. No. 40;		
	Matches	29;	Conservative	10;	Indels 0;
				Gaps	0.
QY	11 CTACGCTTCATCGCTCTCGCGCCGCCGCTGCCCA	49			

11 CTACGCTCATCGCTCTCGTGGCTGCCGCCCTGCCA 49

Db 1025 CGCCCGCTCTTCGCTTCTTCCCTTCCGGGCTCGCA 987

```

RESULT 27
US-09-539-333D-1/c
Sequence 1, Application US/09539333D
Patent No. 6476208
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marla
APPLICANT: Chumakov, Ilya
APPLICANT: Bougueleret, Lydie
APPLICANT: Bihain, Bernard
APPLICANT: Essioux, Laurent
TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
FILE REFERENCE: GENSET.047AUS
CURRENT APPLICATION NUMBER: US/09/539,333D
PRIORITY FILING DATE: 2000-03-10
PRIORITY APPLICATION NUMBER: US 60/126,903
PRIORITY FILING DATE: 1999-03-30
PRIORITY APPLICATION NUMBER: US 60/131,971
PRIORITY FILING DATE: 1999-04-30
PRIORITY APPLICATION NUMBER: US 60/132,065
PRIORITY FILING DATE: 1999-04-30
PRIORITY APPLICATION NUMBER: US 60/143,928
PRIORITY FILING DATE: 1999-07-14
PRIORITY APPLICATION NUMBER: US 60/145,915
PRIORITY FILING DATE: 1999-07-27
PRIORITY APPLICATION NUMBER: US 60/146,453
PRIORITY FILING DATE: 1999-07-29
PRIORITY APPLICATION NUMBER: US 60/146,452
PRIORITY FILING DATE: 1999-07-29
PRIORITY APPLICATION NUMBER: US 60/162,288
PRIORITY FILING DATE: 1999-10-28
PRIORITY APPLICATION NUMBER: US 09/416,384
PRIORITY FILING DATE: 1999-10-12
NUMBER OF SEQ ID NOS: 231
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 319608
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 31..1107
OTHER INFORMATION: 5'regulatory region g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 1108..1289
OTHER INFORMATION: exon A g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 14877..14920
OTHER INFORMATION: exon B g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 18778..18862
OTHER INFORMATION: exon Bbis g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 25593..25740
OTHER INFORMATION: exon C g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 29388..29502
OTHER INFORMATION: exon D g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 29967..30282
OTHER INFORMATION: exon E g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 64666..64812

```

```

1 OTHER INFORMATION: exon F g35018 gene
2 FEATURE:
3 NAME/KEY: exon
4 LOCATION: 65505..65853
5 OTHER INFORMATION: exon G g35018 gene
6 FEATURE:
7 NAME/KEY: misc feature
8 LOCATION: 65854..67854
9 OTHER INFORMATION: 3' regulatory region g35018 gene
10 FEATURE:
11 NAME/KEY: exon
12 LOCATION: 94124..94964
13 OTHER INFORMATION: exon g35017
14 FEATURE:
15 NAME/KEY: exon
16 LOCATION: 201188..201234
17 OTHER INFORMATION: exon S g35030 gene
18 FEATURE:
19 NAME/KEY: exon
20 LOCATION: 214676..214793
21 OTHER INFORMATION: exon T g35030 gene
22 FEATURE:
23 NAME/KEY: exon
24 LOCATION: 215702..215746
25 OTHER INFORMATION: exon U g35030 gene
26 FEATURE:
27 NAME/KEY: exon
28 LOCATION: 216836..216915
29 OTHER INFORMATION: exon V g35030 gene
30 FEATURE:
31 NAME/KEY: misc feature
32 LOCATION: 213818..215818
33 OTHER INFORMATION: 3' regulatory region g34872 gene
34 FEATURE:
35 NAME/KEY: exon
36 LOCATION: 215819..215941
37 OTHER INFORMATION: exon R complement g34872 gene
38 FEATURE:
39 NAME/KEY: exon
40 LOCATION: 215819..215975
41 OTHER INFORMATION: exon Rbis complement g34872 gene
42 FEATURE:
43 NAME/KEY: exon
44 LOCATION: 216661..216952
45 OTHER INFORMATION: exon Qbis complement g34872 gene
46 FEATURE:
47 NAME/KEY: exon
48 LOCATION: 216661..217061
49 OTHER INFORMATION: exon Q complement g34872 gene
50 FEATURE:
51 NAME/KEY: exon
52 LOCATION: 217027..217061
53 OTHER INFORMATION: exon Q1 complement g34872 gene
54 FEATURE:
55 NAME/KEY: exon
56 LOCATION: 228647..229742
57 OTHER INFORMATION: exon X complement g34872 gene
58 FEATURE:
59 NAME/KEY: exon
60 LOCATION: 230408..230721
61 OTHER INFORMATION: exon P complement g34872 gene
62 FEATURE:
63 NAME/KEY: exon
64 LOCATION: 231272..231412
65 OTHER INFORMATION: exon Obis complement g34872 gene
66 FEATURE:
67 NAME/KEY: exon
68 LOCATION: 231787..231880
69 OTHER INFORMATION: exon O2 complement g34872 gene
70 FEATURE:
71 NAME/KEY: exon
72 LOCATION: 231870..231879
73 OTHER INFORMATION: exon O1 complement g34872 gene

```

```

1 FEATURE:
2 NAME/KEY: exon
3 LOCATION: 234174..234321
4 OTHER INFORMATION: exon O complement g34872 gene
5 FEATURE:
6 NAME/KEY: exon
7 LOCATION: 237406..237428
8 OTHER INFORMATION: exon Nbis complement g34872 gene
9 FEATURE:
10 NAME/KEY: exon
11 LOCATION: 239719..239807
12 OTHER INFORMATION: exon N2 complement g34872 gene
13 FEATURE:
14 NAME/KEY: exon
15 LOCATION: 239719..239853
16 OTHER INFORMATION: exon N complement g34872 gene
17 FEATURE:
18 NAME/KEY: exon
19 LOCATION: 240528..240569
20 OTHER INFORMATION: exon M117 complement g34872 gene
21 FEATURE:
22 NAME/KEY: exon
23 LOCATION: 240528..240596
24 OTHER INFORMATION: exon M1090 complement g34872 gene
25 FEATURE:
26 NAME/KEY: exon
27 LOCATION: 240528..240617
28 OTHER INFORMATION: exon M1069 complement g34872 gene
29 FEATURE:
30 NAME/KEY: exon
31 LOCATION: 240528..240644
32 OTHER INFORMATION: exon MS2 complement g34872 gene
33 FEATURE:
34 NAME/KEY: exon
35 LOCATION: 240528..240824
36 OTHER INFORMATION: exon M862 complement g34872 gene
37 FEATURE:
38 NAME/KEY: exon
39 LOCATION: 240528..240994
40 OTHER INFORMATION: exon M692 complement g34872 gene
41 FEATURE:
42 NAME/KEY: exon
43 LOCATION: 240528..241685
44 OTHER INFORMATION: exon M1 complement g34872 gene
45 FEATURE:
46 NAME/KEY: exon
47 LOCATION: 240800..240993
48 OTHER INFORMATION: exon MS1 complement g34872 gene
49 FEATURE:
50 NAME/KEY: misc feature
51 LOCATION: 241686..243685
52 OTHER INFORMATION: 5.regulatory region g34872 gene
53 FEATURE:
54 NAME/KEY: misc feature
55 LOCATION: 290652..297652
56 OTHER INFORMATION: 3.regulatory region g34665 gene
57 FEATURE:
58 NAME/KEY: exon
59 LOCATION: 292653..292841

```

```

Query Match      44.2%  Score 23  DB 4;  Length 319608;
Best Local Similarity 68.1%  Pred. No. 65;
Matches 32;  Conservative 0;  Mismatches 15;  Indels 0;  Gaps 0;

Oy      4  CAGTTCCTCAGCGCTCATCGCTCTGCTGCGTCCGCGCTCGGCAG 50
      |||||  |||||  |||||  |||||  |||||  |||||
Db      184156  CAGTTCCTCAGCGCATCGCTCTGCTGCTGTCATATCCAG 184110

RESULT 28
US-09-679-409-1/C
; Sequence 1, Application US/09679409
; Patent No. 6555316

```

```
/ GENERAL INFORMATION:
/ APPLICANT: Cohen, Daniel
/ APPLICANT: Blumenfeld, Marla
/ APPLICANT: Chumakov, Ilya
/ APPLICANT: Bougueleret, Lydie
/ APPLICANT: Essioux, Laurent
/ TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENE, PROTEINS AND BIALLELIC MARKERS
/ FILE REFERENCE: 53.US15.CIP
/ CURRENT APPLICATION NUMBER: US/09/679,409
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 09/539,333
/ PRIOR FILING DATE: 2000-03-03
/ PRIOR APPLICATION NUMBER: 09/416,384
/ PRIOR FILING DATE: 1999-10-12
/ PRIOR APPLICATION NUMBER: 60/168,088
/ PRIOR FILING DATE: 1999-11-30
/ NUMBER OF SEQ. ID NOS: 134
/ SOFTWARE: Patent.pm
/ SEQ ID NO 1
/ LENGTH: 319608
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: 199122..201122
/ OTHER INFORMATION: 5'regulatory region
/ NAME/KEY: exon
/ LOCATION: 201123..201234
/ OTHER INFORMATION: exon S
/ NAME/KEY: exon
/ LOCATION: 201123..201560
/ OTHER INFORMATION: exon S2
/ NAME/KEY: exon
/ LOCATION: 214676..214793
/ OTHER INFORMATION: exon T
/ NAME/KEY: exon
/ LOCATION: 215702..215746
/ OTHER INFORMATION: exon U
/ NAME/KEY: exon
/ LOCATION: 216836..216994
/ OTHER INFORMATION: exon V
/ NAME/KEY: exon
/ LOCATION: 216836..217077
/ OTHER INFORMATION: exon V2
/ NAME/KEY: exon
/ LOCATION: 217671..217764
/ OTHER INFORMATION: exon V1
/ NAME/KEY: exon
/ LOCATION: 227655..227736
/ OTHER INFORMATION: exon V4
/ NAME/KEY: exon
/ LOCATION: 238715..238919
/ OTHER INFORMATION: exon V3
/ NAME/KEY: exon
/ LOCATION: 240440..240673
/ OTHER INFORMATION: exon W
/ NAME/KEY: exon
/ LOCATION: 240440..241153
/ OTHER INFORMATION: exon W2
/ NAME/KEY: exon
/ LOCATION: 241072..241291
/ OTHER INFORMATION: exon X
/ NAME/KEY: exon
/ LOCATION: 244353..244561
/ OTHER INFORMATION: exon Y
/ NAME/KEY: exon
/ LOCATION: 246273..247802
/ OTHER INFORMATION: exon Z
/ NAME/KEY: misc.feature
/ LOCATION: 247803..249803
/ OTHER INFORMATION: 3'regulatory region
/ NAME/KEY: allele
/ LOCATION: 8316

/ OTHER INFORMATION: 99-27943-150 : polymorphic base G or C
/ NAME/KEY: allele
/ LOCATION: 21672
/ OTHER INFORMATION: 99-27935-193 : polymorphic base G or C
/ NAME/KEY: allele
/ LOCATION: 65485
/ OTHER INFORMATION: 8-128-33 : polymorphic base C or T
/ NAME/KEY: allele
/ LOCATION: 95396
/ OTHER INFORMATION: 99-31960-363 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 107281
/ OTHER INFORMATION: 99-24656-260 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 160640
/ OTHER INFORMATION: 99-24639-163 : polymorphic base A or C
/ NAME/KEY: allele
/ LOCATION: 160876
/ OTHER INFORMATION: 99-24634-108 : polymorphic base A or T
/ NAME/KEY: allele
/ LOCATION: 168974
/ OTHER INFORMATION: 99-7652-162 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 170810
/ OTHER INFORMATION: 99-16100-147 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 173358
/ OTHER INFORMATION: 99-5862-167 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 189957
/ OTHER INFORMATION: 99-5919-215 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 197163
/ OTHER INFORMATION: 99-24658-410 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 200778
/ OTHER INFORMATION: 8-303-235 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 202651
/ OTHER INFORMATION: 8-300-221 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 202679
/ OTHER INFORMATION: 8-300-193 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 203378
/ OTHER INFORMATION: 8-299-128 : polymorphic base A or T
/ NAME/KEY: allele
/ LOCATION: 204138
/ OTHER INFORMATION: 8-296-213 : polymorphic base A or T
/ NAME/KEY: allele
/ LOCATION: 204605
/ OTHER INFORMATION: 8-252-190 : polymorphic base C or T
/ NAME/KEY: allele
/ LOCATION: 204934
/ OTHER INFORMATION: 99-24644-194 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 205206
/ OTHER INFORMATION: 8-295-248 : polymorphic base A or C
/ NAME/KEY: allele
/ LOCATION: 205329
/ OTHER INFORMATION: 8-295-125 : polymorphic base C or T
/ NAME/KEY: allele
/ LOCATION: 206064
/ OTHER INFORMATION: 8-293-130 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 206545
/ OTHER INFORMATION: 8-292-198 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 207313
/ OTHER INFORMATION: 8-251-322 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 208285
/ OTHER INFORMATION: 8-289-322 : polymorphic base A or G
```



```
APPLICATION NUMBER: 08/169,948
FILING DATE: 17-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Stone, Christopher L.
REGISTRATION NUMBER: 35,696
REFERENCE/DOCKET NUMBER: GC226D14
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 742-7555
TELEFAX: (415) 742-7217
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-448-873-35

Query Match      43.1%; Score 22.4; DB 2; Length 48;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY      1 ATGCACTTCTTCACGCTCTCATGCTGCTGCGGCGCTCGCC 48
Db      1 ATGAAGTTCCTTAAGTCTCTCCCTGCTCATACCGCGCCCTTGCC 48

RESULT 31
US-08-382-452D-35
Sequence 35, Application US/08382452D
Patent No. 6268196
GENERAL INFORMATION:
APPLICANT: Fowler, Timothy
APPLICANT: Clarkson, Kathleen A.
APPLICANT: Ward, Michael
APPLICANT: Collier, Katherine D.
APPLICANT: Larenas, Edmund A.
TITLE OF INVENTION: NOVEL CELLULOSE ENZYMES AND SYSTEMS
TITLE OF INVENTION: FOR THEIR EXPRESSION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/382,452D
FILING DATE: February 1, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Christopher L. Stone
REGISTRATION NUMBER: 36,696
REFERENCE/DOCKET NUMBER: GC226-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 742-7555
TELEFAX: (415) 742-7217
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-382-452D-35

Query Match      43.1%; Score 22.4; DB 3; Length 48;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

```
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY      1 ATGCACTTCTTCACGCTCTCATGCTGCTGCGGCGCTCGCC 48
Db      1 ATGAAGTTCCTTAAGTCTCTCCCTGCTCATACCGCGCCCTTGCC 48

RESULT 32
US-08-507-362A-19
Sequence 19, Application US/08507362A
Patent No. 6562340
GENERAL INFORMATION:
APPLICANT: Bedford, Michael
APPLICANT: Morgan, Andrew
APPLICANT: Fowler, Timothy
APPLICANT: Ward, Michael
APPLICANT: Clarkson, Kathleen
APPLICANT: Collier, Katherine
APPLICANT: Larenas, Edmund
TITLE OF INVENTION: An Enzyme Feed Additive and Animal Feed Including It
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/507,362A
FILING DATE: 27-Oct-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Castaneda, Janet
REGISTRATION NUMBER: 33,228
REFERENCE/DOCKET NUMBER: GC226-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 846-4072
TELEFAX: (650) 845-6504
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-08-507-362A-19

Query Match      43.1%; Score 22.4; DB 4; Length 49;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY      1 ATGCACTTCTTCACGCTCTCATGCTGCTGCGGCGCTCGCC 48
Db      1 ATGAAGTTCCTTAAGTCTCTCCCTGCTCATACCGCGCCCTTGCC 48

RESULT 33
US-09-146-770-2
Sequence 2, Application US/09146770
Patent No. 6187732
GENERAL INFORMATION:
APPLICANT: Fowler, Timothy
APPLICANT: Mutant Egit Cellulase, DNA Encoding
TITLE OF INVENTION: Such Egit Cellulase Compositions and Methods for Obtaining Same
FILE REFERENCE: GC546
CURRENT APPLICATION NUMBER: US/09/146,770
```


ADDRESSEE: Genencor International
STREET: 180 Kimball Way
CITY: South San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/032,848C
FILING DATE: MAR 17 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Horn, Margaret A.
REGISTRATION NUMBER: 33,401
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 742-7356
TELEFAX: 415 742-7217
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1050 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-032-848C-11

Query Match 43.1%; Score 22.4; DB 1; Length 1050;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGCACTCTCTCAGCGCTCTCATCGCTCTGCTGCGCGCCCTCGCC 48
DB 151 ATGAAGTCTCTCAAGTCTCTCCCTGCGCTCATACCGCGCCCTGCGC 198

RESULT 39
US-08-438-870-11
Sequence 11, Application US/08438870
Patent No. 5753484
GENERAL INFORMATION:
APPLICANT: Ward, Michael
APPLICANT: Clarkson, Kathleen A.
APPLICANT: Weiss, Geoffrey L.
APPLICANT: Larenas, Edward
APPLICANT: Lorch, Jeffrey D.
TITLE OF INVENTION: Purification and Molecular Cloning of EG
TITLE OF INVENTION: 111 Cellulase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genencor International
STREET: 180 Kimball Way
CITY: South San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/438,870
FILING DATE: MAY 10, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Christopher L. Stone
REGISTRATION NUMBER: 35,696
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 742-7555

TELEFAX: 415 742-7217
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1050 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-438-870-11

Query Match 43.1%; Score 22.4; DB 1; Length 1050;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGCACTCTCTCAGCGCTCTCATCGCTCTGCTGCGCGCCCTCGCC 48
DB 151 ATGAAGTCTCTCAAGTCTCTCCCTGCGCTCATACCGCGCCCTGCGC 198

RESULT 40
US-08-774-065-1
Sequence 1, Application US/08774065
Patent No. 5969899
GENERAL INFORMATION:
APPLICANT: Bower, Benjamin
APPLICANT: Clarkson, Kathleen
APPLICANT: Larenas, Edmund
TITLE OF INVENTION: NOVEL OVERSIZED CELLULASE COMPOSITIONS
TITLE OF INVENTION: FOR USE IN DIFFERENT COMPOSITIONS AND
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: GENENCOR INTERNATIONAL
STREET: 925 PAGE MILL ROAD
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/774,065
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Glaister, Debra J.
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC368
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-846-7620
TELEFAX: 415-845-6504
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1050 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-774-065-1

Query Match 43.1%; Score 22.4; DB 2; Length 1050;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGCACTCTCTCAGCGCTCTCATCGCTCTGCTGCGCGCCCTCGCC 48
DB 151 ATGAAGTCTCTCAAGTCTCTCCCTGCGCTCATACCGCGCCCTGCGC 198

Fri Nov 21 10:36:28 2003

us-10-081-935-1.rni

Page 17

Search completed: November 13, 2003, 08:53:19
Job time : 65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2003, 08:18:16 (Search time 216 Seconds

(without alignments)
786.790 Million cell updates/sec

Title: US-10-081-935-1

Perfect score: 52
Sequence: 1 atgcagctctctcacgctctc.....cgctgcgagcgcctgcagctg 52

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	48.1	336	9 US-09-825-414-37	Sequence 37, Appl
2	25	48.1	1920	9 US-09-815-242-7958	Sequence 7958, Ap
3	24.6	47.3	1566	9 US-09-804-551B-43	Sequence 43, Appl
4	24.6	47.3	1566	14 US-10-270-333-149	Sequence 149, App
5	24.6	47.3	3772	14 US-10-270-333-148	Sequence 148, App
6	24.4	46.9	925	12 US-10-029-386-7131	Sequence 7131, Ap
7	24.4	46.9	912	14 US-10-156-761-6808	Sequence 6808, Ap
8	24.4	46.9	954	14 US-10-156-761-7129	Sequence 7129, Ap
9	24.4	46.9	9025608	14 US-10-156-761-1	Sequence 1, Appl
10	22.8	45.8	609	10 US-09-738-626-3337	Sequence 3337, Ap
11	23.8	45.8	1455	14 US-10-156-761-2054	Sequence 2054, Ap
12	23.8	45.8	3309400	10 US-09-738-626-1	Sequence 1, Appl
13	23.6	45.4	590	12 US-10-027-632-50396	Sequence 50396, A
14	23.6	45.4	590	12 US-10-027-632-50397	Sequence 50397, A
15	23.6	45.4	590	12 US-10-027-632-50398	Sequence 50398, A
16	23.6	45.4	590	12 US-10-027-632-82972	Sequence 82972, A

17	23.6	45.4	590	12 US-10-027-632-82973	Sequence 82973, A
18	23.6	45.4	590	12 US-10-027-632-82974	Sequence 82974, A
19	23.6	45.4	590	13 US-10-027-632-50396	Sequence 50396, A
20	23.6	45.4	590	13 US-10-027-632-50397	Sequence 50397, A
21	23.6	45.4	590	13 US-10-027-632-50398	Sequence 50398, A
22	23.6	45.4	590	13 US-10-027-632-82972	Sequence 82972, A
23	23.6	45.4	590	13 US-10-027-632-82973	Sequence 82973, A
24	23.6	45.4	590	13 US-10-027-632-82974	Sequence 82974, A
25	23.6	45.4	1136	14 US-10-198-846-12577	Sequence 12577, A
26	23.6	45.4	135638	12 US-10-314-657-1	Sequence 1, Appl1
27	23.4	45.0	1029	14 US-10-156-761-5574	Sequence 5574, Ap
28	23.4	45.0	2466	12 US-09-953-348-83	Sequence 83, Appl1
29	23.4	45.0	2466	14 US-10-267-255-83	Sequence 83, Appl1
30	23.4	45.0	18034	12 US-09-953-348-75	Sequence 75, Appl1
31	23.4	45.0	18034	14 US-10-267-255-75	Sequence 75, Appl1
32	23.4	45.0	9025608	14 US-10-267-255-75	Sequence 1, Appl1
33	23.2	44.6	271	10 US-10-156-761-1	Sequence 12667, A
34	23.2	44.6	450	9 US-09-878-574-12667	Sequence 223, App
35	23.2	44.6	497	12 US-09-667-550-223	Sequence 223, App
36	23.2	44.6	597	12 US-10-029-386-21891	Sequence 8197, Ap
37	23	44.2	492	11 US-09-918-995-37972	Sequence 37972, A
38	23	44.2	939	14 US-10-156-761-4691	Sequence 4691, Ap
39	22.8	43.8	322	10 US-09-960-352-476	Sequence 476, App
40	22.8	43.8	379	10 US-09-960-352-11092	Sequence 11092, A
41	22.8	43.8	381	10 US-09-960-352-12293	Sequence 12293, A
42	22.8	43.8	2559	14 US-10-128-714-7006	Sequence 7006, Ap
43	22.8	43.8	2621	14 US-10-128-714-6006	Sequence 6006, Ap
44	22.8	43.8	3608	14 US-10-128-714-6	Sequence 6, Appl1
45	22.8	43.8	4621	14 US-10-128-714-5006	Sequence 5006, Ap

ALIGNMENTS

RESULT 1
US-09-825-414-37
; Sequence 37, Application US/09825414
; Patent No. US20020083489A1
; GENERAL INFORMATION:
; APPLICANT: Colimer, Alan
; APPLICANT: Alfano, James R.
; APPLICANT: Charkowski, Amy O.
; TITLE OF INVENTION: DNA MOLECULES AND POLYPEPTIDES OF PSEUDOMONAS SYRINGAE
; FILE REFERENCE: 19603/3243
; CURRENT APPLICATION NUMBER: US/09/825,414
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/194,160
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 60/224,604
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/249,548
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 37
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Pseudomonas syringae
; US-09-825-414-37
Query Match 48.1%; Score 25; DB 9; Length 336;
Best Local Similarity 69.4%; Pred. No. 12;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Caps 0;
QY 4 CAGTTCCTCAGCGCTCTCATCGCTCGTGCAGCGCTGCAGT 52
DB 154 CATTGGCGTGGCGCTCTCAACCCCTTGTGATGATCCGCCGCCGATTG 202
RESULT 2
US-09-815-242-7958/c
; Sequence 7958, Application US/09815242

```
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7958
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1920)
US-09-815-242-7958

Query Match          48.1%; Score 25; DB 9; Length 1920;
Best Local Similarity 69.4%; Pred. No. 12;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      4  CAGTTCTCTCAGCCTCTCATCGCTCTGCGCGCGCGCTCGCCACTG 52
DB      507  CAGCTCGCTGACCTCTCTGCTGCTGCGCAGCGCGCAGCTCAGCG 459

RESULT 3
US-09-804-551B-43/C
; Sequence 43, Application US/09804551B
; Patent No. US20020056151A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Receptors for peptides from insects
; FILE REFERENCE: Le A 34 394
; CURRENT APPLICATION NUMBER: US/09/804,551B
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: DE 100 13 618.4
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1563)
US-09-804-551B-43

Query Match          47.3%; Score 24.6; DB 9; Length 1566;
```

```
Best Local Similarity 70.2%; Pred. No. 16;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      3  GCAGTTCTCTCAGCCTCTCATCGCTCTGCGCGCGCGCTCGCCCA 49
DB      646  GCACAGCGGTACGACCTCTCTGCTCTCTCTCGCTGAGTCTCCACA 600

RESULT 4
US-10-270-333-149/C
; Sequence 149, Application US/10270333
; Publication No. US20030092124A1
; GENERAL INFORMATION:
; APPLICANT: Cravchik, Anibal
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES
; FILE REFERENCE: CL000733CON
; CURRENT APPLICATION NUMBER: US/10/270,333
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/168,677
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/175,691
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/191,638
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 149
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Drosophila
US-10-270-333-149

Query Match          47.3%; Score 24.6; DB 14; Length 1566;
Best Local Similarity 70.2%; Pred. No. 16;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      3  GCAGTTCTCTCAGCCTCTCATCGCTCTGCGCGCGCGCTCGCCCA 49
DB      646  GCACAGCGGTACGACCTCTCTGCTCTCTCTCGCTGAGTCTCCACA 600

RESULT 5
US-10-270-333-148/C
; Sequence 148, Application US/10270333
; Publication No. US20030092124A1
; GENERAL INFORMATION:
; APPLICANT: Cravchik, Anibal
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES
; FILE REFERENCE: CL000733CON
; CURRENT APPLICATION NUMBER: US/10/270,333
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/168,677
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/175,691
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/191,638
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148
; LENGTH: 3772
; TYPE: DNA
; ORGANISM: Drosophila
US-10-270-333-148

Query Match          47.3%; Score 24.6; DB 14; Length 3772;
Best Local Similarity 70.2%; Pred. No. 16;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```

; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA

```

```

; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

```

```

Query Match          46.9%; Score 24.4; DB 14; Length 9025608;
Best Local Similarity 68.0%; Pred. No. 14;
Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

```

QY      3 GCAGTCTCTCAGCCTCTCATCGCTCTCGTCGCGCGGCTGGCCAGTG 52
DB      8154610 GCAGGTGGCGCAGCGGCTCGACGACGTTGCCGCTGCCGCTCTTG 8154561

```

```

RESULT 10
US-09-738-626-3337
; Sequence 3337, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3337
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3337

```

```

Query Match          45.8%; Score 23.8; DB 10; Length 609;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

```

QY      2 TGCAGTTCTCTCAGCCTCTCATCGCTCTCGTCGCGCGGCTGGCCAGTG 52
DB      401 TTCAGATCGGATACGCCATCATCGGTATCCGCCGAGAGGTCTCGCCGCTG 451

```

```

RESULT 11
US-10-156-761-2054
; Sequence 2054, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHITA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761

```

```

; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 2054
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1455)
US-10-156-761-2054

```

```

Query Match          45.8%; Score 23.8; DB 14; Length 1455;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

```

QY      1 ATGCAGTTCTCTCAGCCTCTCATCGCTCTCGTCGCGCGGCTGGCCAGTG 51
DB      1282 ATGCCGTTCCAGTACATCATCACCGCGGCTGGCGGCTGCTCACT 1332

```

```

RESULT 12
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

```

```

Query Match          45.8%; Score 23.8; DB 10; Length 3309400;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

```

QY      2 TGCAGTTCTCTCAGCCTCTCATCGCTCTCGTCGCGCGGCTGGCCAGTG 52
DB      3218393 TTCAGATCGGATACGCCATCATCGGTATCCGCCGAGAGGTCTCGCCGCTG 3218443

```

```

RESULT 13
US-10-027-632-50396/c
; Sequence 50396, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

```

```
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 50396
/ LENGTH: 590
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-50396
```

```
Query Match          45.4%; Score 23.6; DB 12; Length 590;
Best Local Similarity 76.3%; Pred. No. 38;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY      2 TGCAGTTCCTCAGCCTCTCATCGCTCTCGTCGCTGCC 39
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      264 TCGTGTTCCTCAGCCACACACACTCTCTGCTCCCTACC 227
```

```
RESULT 14
/ Sequence 50397, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 50397
/ LENGTH: 590
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-50397
```

```
Query Match          45.4%; Score 23.6; DB 12; Length 590;
Best Local Similarity 76.3%; Pred. No. 38;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY      2 TGCAGTTCCTCAGCCTCTCATCGCTCTCTGCTGCC 39
```

```
Db      264 TCGTGTTCCTCAGCCACACACACTCTCTGCTCCCTACC 227
```

```
RESULT 15
/ Sequence 50398, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 50398
/ LENGTH: 590
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-50398
```

```
Query Match          45.4%; Score 23.6; DB 12; Length 590;
Best Local Similarity 76.3%; Pred. No. 38;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY      2 TGCAGTTCCTCAGCCTCTCATCGCTCTCGTCGCTGCC 39
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      264 TCGTGTTCCTCAGCCACACACACTCTCTGCTCCCTACC 227
```

```
RESULT 16
/ Sequence 82972, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 82972
LENGTH: 590
TYPE: DNA
ORGANISM: Human
US-10-027-632-82972

Query Match 45.4%; Score 23.6; DB 12; Length 590;
Best Local Similarity 76.3%; Pred. No. 38;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 TGCAGTCTCTCAGCGCTCTCATCGCTCTCGTCCGCC 39
DB 327 TCGTGTCTTCTCAAGCCACACACTCTCGTCCCTACC 364

RESULT 17
US-10-027-632-82973

Sequence 82973, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 82973

LENGTH: 590

TYPE: DNA

ORGANISM: Human

US-10-027-632-82973

Query Match 45.4%; Score 23.6; DB 12; Length 590;
Best Local Similarity 76.3%; Pred. No. 38;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 TGCAGTCTCTCAGCGCTCTCATCGCTCTCGTCCGCC 39
DB 327 TCGTGTCTTCTCAAGCCACACACTCTCGTCCCTACC 364

RESULT 18
US-10-027-632-82974

Sequence 82974, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 82974
LENGTH: 590
TYPE: DNA
ORGANISM: Human
US-10-027-632-82974

Query Match 45.4%; Score 23.6; DB 12; Length 590;
Best Local Similarity 76.3%; Pred. No. 38;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 TGCAGTCTCTCAGCGCTCTCATCGCTCTCGTCCGCC 39
DB 327 TCGTGTCTTCTCAAGCCACACACTCTCGTCCCTACC 364

RESULT 19
US-10-027-632-50396/c

Sequence 50396, Application US/10027632
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 50396

LENGTH: 590

TYPE: DNA

ORGANISM: Human

US-10-027-632-50396

Query Match 45.4%; Score 23.6; DB 13; Length 590;
Best Local Similarity 76.3%; Pred. No. 38;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 TGCAGTCTCTCAGCGCTCTCATCGCTCTCGTCCGCC 39
DB 264 TCGTGTCTTCTCAAGCCACACACTCTCGTCCCTACC 227

RESULT 20
US-10-027-632-50397/c

Sequence 50397, Application US/10027632
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

US-10-027-632-50397/c
Sequence 50397, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

ORGANISM: Human
US-10-027-632-82973

Query Match 45.4%; Score 23.6; DB 13; Length 590;
Best Local Similarity 76.3%; Pred. No. 38;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 TGCAGTTCCTCAGCGCTCTCATCGCTCTCGTCGCTGCC 39
DB 327 TGCTGTTTCTCAAGCCACACACTCTCGTCCCTACC 364

RESULT 24
US-10-027-632-82974
Sequence 82974, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIORITY FILING DATE: 2002-04-30
PRIORITY APPLICATION NUMBER: US 60/218,006
PRIORITY FILING DATE: 2000-07-12
PRIORITY APPLICATION NUMBER: US 60/198,676
PRIORITY FILING DATE: 2000-04-20
PRIORITY APPLICATION NUMBER: US 60/193,483
PRIORITY FILING DATE: 2000-03-29
PRIORITY APPLICATION NUMBER: US 60/185,218
PRIORITY FILING DATE: 2000-02-24
PRIORITY APPLICATION NUMBER: US 60/167,363
PRIORITY FILING DATE: 1999-11-23
PRIORITY APPLICATION NUMBER: US 60/156,358
PRIORITY FILING DATE: 1999-09-28
PRIORITY APPLICATION NUMBER: US 60/146,002
PRIORITY FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 82974
LENGTH: 590
TYPE: DNA
ORGANISM: Human
US-10-027-632-82974

Query Match 45.4%; Score 23.6; DB 13; Length 590;
Best Local Similarity 76.3%; Pred. No. 38;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 TGCAGTTCCTCAGCGCTCTCATCGCTCTCGTCGCTGCC 39
DB 327 TGCTGTTTCTCAAGCCACACACTCTCGTCCCTACC 364

RESULT 25
US-10-198-846-12577
Sequence 12577, Application US/10198846
GENERAL INFORMATION:
APPLICANT: Li, James

APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinhilber, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
PRIORITY FILING DATE: 2002-07-18
PRIORITY APPLICATION NUMBER: 60/306,220
PRIORITY FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12577

LENGTH: 1136
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1127, 1128, 1129, 1130, 1131, 1132, 1133, 1134, 1135, 1136
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12577

Query Match 45.4%; Score 23.6; DB 14; Length 1136;
Best Local Similarity 69.6%; Pred. No. 37;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ATGCAGTTCCTCAGCGCTCTCATCGCTCTCGTCGCTGCCGACCTGC 46
DB 54 ATGCAGTTCCTCAGCGCTCTCATCGCTCTCGTCGCTGCCGACCTGC 99

RESULT 26
US-10-314-657-1/C
Sequence 1, Application US/10314657
Publication No. US20030175888A1
GENERAL INFORMATION:
APPLICANT: SHEN, Ben

APPLICANT: TANG, Gong-Li
TITLE OF INVENTION: Discrete Acyltransferases Associated with Type I Polyketide
FILE REFERENCE: 054030-0021
CURRENT APPLICATION NUMBER: US/10/314,657
PRIORITY FILING DATE: 2002-12-09
PRIORITY APPLICATION NUMBER: PCT/US02/08937
PRIORITY FILING DATE: 2002-03-22
PRIORITY APPLICATION NUMBER: US 60/278,935
PRIORITY FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 214
SOFTWARE: PatentIn Version 3.2
SEQ ID NO 1
LENGTH: 135638
TYPE: DNA
ORGANISM: Streptomyces atroolivaceus
US-10-314-657-1

Query Match 45.4%; Score 23.6; DB 12; Length 135638;
Best Local Similarity 76.3%; Pred. No. 32;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 15 CGCTCTCATCGCTCTCATCGCTCTCGTCGCTGCCGACCTGC 52
DB 34257 CGCTCTCATCGCTCTCATCGCTCTCGTCGCTGCCGACCTGC 34220

RESULT 27
US-10-156-761-5574
Sequence 5574, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO
APPLICANT: ISHIDA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
PRIORITY FILING DATE: 2002-05-29
PRIORITY APPLICATION NUMBER: JP 2001-204089
PRIORITY FILING DATE: 2001-05-30
PRIORITY APPLICATION NUMBER: JP 2001-272697
PRIORITY FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109

```

; SEQ ID NO 5574
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1029)
US-10-156-761-5574

```

```

Query Match          45.0%; Score 23.4; DB 14; Length 1029;
Best Local Similarity 67.3%; Pred. No. 43;
Matches 33; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

```

Qy 1 ATGCACTCTCTCAGCGCTCATCGCTCTGCTGCGCGGCTCGCCA 49
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10 AAGCGTGAAGTCTCTCTCTGACCGCCCTGCTCGCGGCTCTCGCCA 58

```

```

RESULT 28
US-09-953-348-83
; Sequence 83, Application US/09953348
; Publication No. US20030134398A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, David. H
; APPLICANT: Mao, Yingqing
; APPLICANT: Varoglu, Mustafa
; APPLICANT: He, Min
; APPLICANT: Sheldon, Paul
; TITLE OF INVENTION: MITOMYCIN BIOSYNTHETIC GENE CLUSTER
; FILE REFERENCE: 600.530US1
; CURRENT APPLICATION NUMBER: US/09/953,348
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: PCT/US00/06394
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/266965
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 2466
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
US-09-953-348-83

```

```

Query Match          45.0%; Score 23.4; DB 12; Length 2466;
Best Local Similarity 73.2%; Pred. No. 42;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

```

Qy 9 CTCACGCTCTCATCGCTCTGCTGCGCGGCTCGCCA 49
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 48 CGCTGCCCGCCTGATGCGCCTGCTGCGCTCGCGCACCGCCA 88

```

```

RESULT 29
US-10-267-255-83
; Sequence 83, Application US/10267255
; Publication No. US20030124689A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600.456US1
; CURRENT APPLICATION NUMBER: US/10/267,255
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 09/266,965
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 08/624,447
; PRIOR FILING DATE: 1996-08-19
; PRIOR APPLICATION NUMBER: PCT/US94/11279
; PRIOR FILING DATE: 1994-10-06

```

```

; PRIOR APPLICATION NUMBER: US 08/133,963
; PRIOR FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 83
; LENGTH: 2466
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
US-10-267-255-83

```

```

Query Match          45.0%; Score 23.4; DB 14; Length 2466;
Best Local Similarity 73.2%; Pred. No. 42;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

```

Qy 9 CTCACGCTCTCATCGCTCTGCTGCGCGGCTCGCCA 49
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 48 CGCTGCCCGCCTGATGCGCCTGCTGCGCTCGCGCACCGCCA 88

```

```

RESULT 30
US-09-953-348-75
; Sequence 75, Application US/09953348
; Publication No. US20030134398A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, David. H
; APPLICANT: Mao, Yingqing
; APPLICANT: Varoglu, Mustafa
; APPLICANT: He, Min
; APPLICANT: Sheldon, Paul
; TITLE OF INVENTION: MITOMYCIN BIOSYNTHETIC GENE CLUSTER
; FILE REFERENCE: 600.530US1
; CURRENT APPLICATION NUMBER: US/09/953,348
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: PCT/US00/06394
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/266965
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 18034
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (302)..(302)
; OTHER INFORMATION: n is a or t or g or c
US-09-953-348-75

```

```

Query Match          45.0%; Score 23.4; DB 12; Length 18034;
Best Local Similarity 73.2%; Pred. No. 40;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

```

Qy 9 CTCACGCTCTCATCGCTCTGCTGCGCGGCTCGCCA 49
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 11562 CGCTGCCCGCCTGATGCGCCTGCTGCGCTCGCGCACCGCCA 11602

```

```

RESULT 31
US-10-267-255-75
; Sequence 75, Application US/10267255
; Publication No. US20030124689A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600.456US1
; CURRENT APPLICATION NUMBER: US/10/267,255
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 09/266,965

```

```

; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 08/624,447
; PRIOR FILING DATE: 1996-08-19
; PRIOR APPLICATION NUMBER: PCT/US94/11279
; PRIOR FILING DATE: 1994-10-06
; PRIOR APPLICATION NUMBER: US 08/133,963
; PRIOR FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 75
; LENGTH: 18034
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (302)...(302)
; OTHER INFORMATION: n is a or t or g or c
US-10-267-255-75
```

```

Query Match          45.0%; Score 23.4; DB 14; Length 18034;
Best Local Similarity 73.2%; Pred. No. 40;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

QY 9 CTCACGCTCTCATGCGCTCTGTCGCGCGGCTTGCCA 49
DB 11562 CGCTGCGCGCCCTGATGCGCTCTGTCGCGGACCGCA 11602
```

```

RESULT 32
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermectilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1
```

```

Query Match          45.0%; Score 23.4; DB 14; Length 9025608;
Best Local Similarity 67.3%; Pred. No. 31;
Matches 33; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

```

QY 1 ATGCAGTTCTCTCAGCGCTCTCATGCGCTCTGTCGCGCGGCTTGCCA 49
DB 6765271 AAGCGTACGTCCTCTCTGACGCGCTCTGTCGCGCGGCTTGCCA 6765319
```

```

RESULT 33
US-09-878-574-12667/c
; Sequence 12667, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Placids
; FILE REFERENCE: 38-21 (15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 12667
; LENGTH: 271
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(271)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 70106061H1
US-09-878-574-12667
```

```

Query Match          44.6%; Score 23.2; DB 10; Length 271;
Best Local Similarity 70.5%; Pred. No. 53;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```

QY 8 TCTTCACGCTCTCATGCGCTCTGTCGCGCGGCTTGCGCAGT 51
DB 59 TCTGCTCTCTCTCATGCGCTCTCATCTCATCTTGTGCGCCAGCAGT 16
```

```

RESULT 34
US-09-867-550-223/c
; Sequence 223, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells an
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 223
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(31)
; OTHER INFORMATION: wherein any n is one of a or t or g or c
US-09-867-550-223
```

```

Query Match          44.6%; Score 23.2; DB 9; Length 450;
Best Local Similarity 77.8%; Pred. No. 52;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```

QY 13 CAGGCTCTCATGCGCTCTGTCGCGCGGCTTGCGCC 48
DB 207 CTCGATCTCGCGGCTCTGTCGCGGCTTCTCTGCCC 172
```

```

RESULT 35
US-10-029-386-21891
; Sequence 21891, Application US/10029386
; Publication No. US20030194704A1
```

```
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: AECOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
NUMBER OF SEQ ID NOS: 34288
SEQ ID NO 21891
LENGTH: 497
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR16.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
OTHER INFORMATION: EST HUMAN HIT: AW295242.1, EVALUE 3.00e-77
OTHER INFORMATION: SWISSPROT HIT: O75351, EVALUE 3.00e-25
US-10-029-386-21891
```

Query Match 44.6%; Score 23.2; DB 12; Length 497;
Best Local Similarity 70.5%; Pred. No. 52;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy 9 CTCACGCTCTCATGCTCTGTCGTCGCCGCGCTCCGCAATG 52
Db 110 CTCGATGCTCTCTCTGCTGCTGTCGTCGCCGCTCATCTGTG 153
```

```
RESULT 36
US-10-029-386-8191
Sequence 8191, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: AECOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 8191
LENGTH: 597
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR16.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
OTHER INFORMATION: EST HUMAN HIT: AW295242.1, EVALUE 0.00e+00
OTHER INFORMATION: EST HUMAN HIT: AW295242.1, EVALUE 3.00e-77
OTHER INFORMATION: SWISSPROT HIT: O75351, EVALUE 9.00e-26
US-10-029-386-8191
```

Query Match 44.6%; Score 23.2; DB 12; Length 597;
Best Local Similarity 70.5%; Pred. No. 52;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy 9 CTCACGCTCTCATGCTCTGTCGTCGCCGCGCTCCGCAATG 52
||| | ||||| | | ||||| ||||| | |||
```

Db 169 CTCGATGCTCTCTGCTGCTGTCGTCGCCGCTCATCTGTG 212

```
RESULT 37
US-09-918-995-37972/c
Sequence 37972, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hysag, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 37972
LENGTH: 492
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(492)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-37972
```

Query Match 44.2%; Score 23; DB 11; Length 492;
Best Local Similarity 68.1%; Pred. No. 61;
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy 4 CAGTTCTTACGCTCTCATGCTCTGTCGTCGCCGCTCCGCAATG 50
Db 134 CGTTCTCTCCCTTCCCTCCCTGCTGCTGTCGTCGCAATGCGCAG 88
```

```
RESULT 38
US-10-156-761-4691
Sequence 4691, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 4691
LENGTH: 939
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(939)
US-10-156-761-4691
```

Query Match 44.2%; Score 23; DB 14; Length 939;
Best Local Similarity 68.1%; Pred. No. 60;
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy 2 TGCAATTCTCTACGCTCTCATGCTCTGTCGTCGCCGCTCCGCAATG 48
||| | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Db 854 TGCAGCGCCCGCCCTTGACGGCTGTCGCCGAGCGCTCTC 900

RESULT 39
US-09-960-352-476
; Sequence 476, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 476
; LENGTH: 322
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 03-BOWMS1-015-Q1-E1-A11
US-09-960-352-476

Query Match 43.8%; Score 22.8; DB 10; Length 322;
Best Local Similarity 71.4%; Pred. No. 73;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 8 TCTCTACGCTCTCATCGCTCTGCTGCGGCGCTCGCCA 49
| | | | | | | | | | | | | | | | | | | | | |
Db 281 TATCGCCCGCTCTTCGCTTCGCTTCCTTCCTTCGCCA 322

RESULT 40
US-09-960-352-11092
; Sequence 11092, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11092
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 48-BOWMS1-002-Q1-E1-D12
US-09-960-352-11092

Query Match 43.8%; Score 22.8; DB 10; Length 379;
Best Local Similarity 71.4%; Pred. No. 73;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 8 TCTCTACGCTCTCATCGCTCTGCTGCGGCGCTCGCCA 49
| | | | | | | | | | | | | | | | | | | | | |
Db 281 TATCGCCCGCTCTTCGCTTCGCTTCCTTCCTTCGCCA 322

Search completed: November 13, 2003, 09:48:44
Job time : 244 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2003, 08:54:45 ; Search time 1969 Seconds

(without alignments)
1080.396 Million cell updates/sec

Title: US-10-081-935-1

Perfect score: 52

Sequence: 1 atgcagctctctcagctctc.....cgctccgcgcctcgagctg 52

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

```

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vtc:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	52	100.0	1201	8	CIU39835
2	52	100.0	1234	8	CIU32518
3	52	100.0	1435	8	CIU51200
4	52	100.0	3821	8	AF013256
5	52	100.0	10822	1	AE004517
6	52	100.0	13170	1	AE0012448
7	52	100.0	75093	1	AF416330
8	52	100.0	134963	1	AP005679
9	52	100.0	146670	8	AP003250
10	52	100.0	156485	2	AP005509
11	52	100.0	157945	2	AP005742
12	52	100.0	172081	2	AP004776
13	52	100.0	298900	1	AP005937
14	52	100.0	1500	1	AF142898
15	52	100.0	2349	8	AY262750
16	52	100.0	63781	2	AC009913
17	52	100.0	64271	2	AC025921
18	52	100.0	75162	2	AC018480
19	52	100.0	99277	2	AC010568
20	52	100.0	101579	2	AC008227
21	52	100.0	117714	3	AC098797
22	52	100.0	123010	2	AC009746
23	52	100.0	124612	2	AC105442
24	52	100.0	127091	2	AC009843
25	52	100.0	137841	2	AC010915
26	52	100.0	149349	8	AC134236
27	52	100.0	157371	3	AC009353
28	52	100.0	159581	2	AP003935
29	52	100.0	160573	9	AC119676
30	52	100.0	161518	2	AC012389
31	52	100.0	173093	2	AC013579
32	52	100.0	173874	3	AE003845
33	52	100.0	185916	3	AC007587
34	52	100.0	187223	2	AC008333
35	52	100.0	188032	2	AC016393
36	52	100.0	188893	2	AC013254
37	52	100.0	188893	2	AC013254
38	52	100.0	188893	2	AC013254
39	52	100.0	188893	2	AC013254
40	52	100.0	188893	2	AC013254
41	52	100.0	188893	2	AC013254
42	52	100.0	188893	2	AC013254
43	52	100.0	188893	2	AC013254
44	52	100.0	188893	2	AC013254
45	52	100.0	188893	2	AC013254
46	52	100.0	188893	2	AC013254
47	52	100.0	188893	2	AC013254
48	52	100.0	188893	2	AC013254
49	52	100.0	188893	2	AC013254
50	52	100.0	188893	2	AC013254
51	52	100.0	188893	2	AC013254
52	52	100.0	188893	2	AC013254
53	52	100.0	188893	2	AC013254
54	52	100.0	188893	2	AC013254
55	52	100.0	188893	2	AC013254
56	52	100.0	188893	2	AC013254
57	52	100.0	188893	2	AC013254
58	52	100.0	188893	2	AC013254
59	52	100.0	188893	2	AC013254
60	52	100.0	188893	2	AC013254
61	52	100.0	188893	2	AC013254
62	52	100.0	188893	2	AC013254
63	52	100.0	188893	2	AC013254
64	52	100.0	188893	2	AC013254
65	52	100.0	188893	2	AC013254

C	66	17	32.7	191022	8	AP003431	AP003431 Oryza sat	C	139	16	30.8	139468	8	AC090441	AC090441 Oryza sat
C	67	17	32.7	200050	1	AL646068	AL646068 Ralscomba	C	140	16	30.8	143618	8	AC026855	AC026855 Homo sapi
C	68	17	32.7	220676	2	AC114574	AC114574 Mus muscu	C	141	16	30.8	150716	8	AC137928	AC137928 Oryza sat
C	69	17	32.7	231461	2	AC096278	AC096278 Rattus no	C	142	16	30.8	152347	2	AC123898	AC123898 Medicago
C	70	17	32.7	244757	3	AE003678	AE003678 Drosophi	C	143	16	30.8	152691	2	AC141704	AC141704 Apis mell
C	71	17	32.7	247590	2	AC119964	AC119964 Mus muscu	C	144	16	30.8	153344	2	AC104160	AC104160 Canis fam
C	72	17	32.7	252602	2	AC091704	AC091704 Mus muscu	C	145	16	30.8	153436	2	AP005907	AP005907 Oryza sat
C	73	17	32.7	263668	2	AC126697	AC126697 Rattus no	C	146	16	30.8	157006	2	AP003860	AP003860 Oryza sat
C	74	17	32.7	299850	1	AP005949	AP005949 Bradyrhiz	C	147	16	30.8	159029	2	AP004692	AP004692 Oryza sat
C	75	16	30.8	550	8	HSA326699	HSA326699 Homo sapi	C	148	16	30.8	160035	9	AC009068	AC009068 Homo sapi
C	76	16	30.8	610	8	AF297566	AF297566 Pelotum	C	149	16	30.8	160743	9	HSK258F22	HSK258F22 Homo sapi
C	77	16	30.8	664	3	DME426941	DME426941 Drosophi	C	150	16	30.8	162339	8	AC022457	AC022457 Oryza sat
C	78	16	30.8	909	1	PSEMVAB	M24016 P. mevaloni	C	151	16	30.8	163278	8	AC008336	AC008336 Drosophi
C	79	16	30.8	1117	1	PSEMVAB	M31807 P. mevaloni	C	152	16	30.8	163450	8	OSJ00127	OSJ00127 Oryza sat
C	80	16	30.8	1286	9	BC000207	BC000207 Homo sapi	C	153	16	30.8	167687	2	AP004001	AP004001 Oryza sat
C	81	16	30.8	1307	9	BC014040	BC014040 Homo sapi	C	154	16	30.8	169006	2	AC142178	AC142178 Rattus no
C	82	16	30.8	1653	6	AX358336	AX358336 Sequence	C	155	16	30.8	169931	2	AC008321	AC008321 Homo sapi
C	83	16	30.8	1705	1	AB029936	AB029936 ButKnoide	C	156	16	30.8	172378	2	BX470161	BX470161 Danio rer
C	84	16	30.8	1871	9	BC043607	BC043607 Homo sapi	C	157	16	30.8	174583	10	AL772313	AL772313 Mouse DNA
C	85	16	30.8	1889	9	BC027475	BC027475 Homo sapi	C	158	16	30.8	176872	10	AL772313	AL772313 Mouse DNA
C	86	16	30.8	1893	3	AF403542	AF403542 Bombyx mo	C	159	16	30.8	177163	9	AC100786	AC100786 Homo sapi
C	87	16	30.8	2248	3	DROEX02	L07660 Drosophila	C	160	16	30.8	177472	9	AC008786	AC008786 Homo sapi
C	88	16	30.8	2484	9	BC015581	BC015581 Homo sapi	C	161	16	30.8	179369	2	AC122291	AC122291 Rattus no
C	89	16	30.8	2585	9	BC007982	BC007982 Homo sapi	C	162	16	30.8	180153	2	AC018798	AC018798 Homo sapi
C	90	16	30.8	2604	3	AF119623	AF119623 Drosophi	C	163	16	30.8	181401	2	AL954127	AL954127 Mus muscu
C	91	16	30.8	2760	1	SRENSRS	Y14336 Streptomyce	C	164	16	30.8	182769	9	AP003176	AP003176 Homo sapi
C	92	16	30.8	3040	3	AF419622	AF419622 Drosophi	C	165	16	30.8	183065	2	AC008401	AC008401 Homo sapi
C	93	16	30.8	3750	1	U93704	U93704 Riffia pach	C	166	16	30.8	183426	2	AC107196	AC107196 Rattus no
C	94	16	30.8	4390	9	BC037556	BC037556 Homo sapi	C	167	16	30.8	184554	3	AC099022	AC099022 Drosophi
C	95	16	30.8	5178	3	AY119603	AY119603 Drosophi	C	168	16	30.8	185200	3	AC092401	AC092401 Drosophi
C	96	16	30.8	7133	9	D80004	D80004 Homo sapien	C	169	16	30.8	185290	2	AC107006	AC107006 Rattus no
C	97	16	30.8	9787	1	AE011672	AE011672 Xanthomon	C	170	16	30.8	185321	9	AC123908	AC123908 Homo sapi
C	98	16	30.8	10049	1	AE014662	AE014662 Bifidobac	C	171	16	30.8	188624	10	AL772282	AL772282 Mouse DNA
C	99	16	30.8	10417	1	AE004546	AE004546 Pseudomob	C	172	16	30.8	188646	9	AC027045	AC027045 Homo sapi
C	100	16	30.8	10444	14	PV1422133	PV1422133 Pseudococ	C	173	16	30.8	189252	2	AC018930	AC018930 Homo sapi
C	101	16	30.8	10542	1	AE001951	AE001951 Rhizobium	C	174	16	30.8	190187	2	AC121140	AC121140 Mus muscu
C	102	16	30.8	10600	1	RM106XOB	AE004805 Pseudomon	C	175	16	30.8	190615	2	AC108411	AC108411 Homo sapi
C	103	16	30.8	10637	1	AE004805	AE007262 Stenohizo	C	176	16	30.8	191162	3	AC008865	AC008865 Homo sapi
C	104	16	30.8	10967	1	AE005841	AE005841 Caulobac	C	177	16	30.8	191558	3	AC013431	AC013431 Drosophi
C	105	16	30.8	11744	4	AY102170	AY102170 Canis fam	C	178	16	30.8	193360	2	AC021977	AC021977 Homo sapi
C	106	16	30.8	13947	4	PSNARXL	Y15552 Pseudomonas	C	179	16	30.8	198677	1	AE001863	AE001863 Homo sapi
C	107	16	30.8	15168	5	AF090339	AF090339 Rhea amer	C	180	16	30.8	198764	9	AC103809	AC103809 Homo sapi
C	108	16	30.8	25108	2	AC140885	AC140885 Botryllus	C	181	16	30.8	200900	9	AC013410	AC013410 Homo sapi
C	109	16	30.8	27655	2	AC019935	AC019935 Drosophi	C	182	16	30.8	201785	2	AC141171	AC141171 Rattus no
C	110	16	30.8	39089	4	AF483210	AF483210 Canis fam	C	183	16	30.8	202190	2	AC026813	AC026813 Mus muscu
C	111	16	30.8	43516	3	AC058781	AC058781 Leishmani	C	184	16	30.8	204712	2	AC134627	AC134627 Rattus no
C	112	16	30.8	49806	9	AC069162	AC069162 Homo sapi	C	185	16	30.8	206387	2	AC130934	AC130934 Homo sapi
C	113	16	30.8	53339	2	AC091847	AC091847 Homo sapi	C	186	16	30.8	209740	10	AL772166	AL772166 Mouse DNA
C	114	16	30.8	58343	1	AB089954	AB089954 Micromono	C	187	16	30.8	214416	2	AC112406	AC112406 Rattus no
C	115	16	30.8	61174	2	AC080104	AC080104 Homo sapi	C	188	16	30.8	216807	2	AC124146	AC124146 Rattus no
C	116	16	30.8	72899	2	AC139009	AC139009 Homo sapi	C	189	16	30.8	219694	2	AC111438	AC111438 Rattus no
C	117	16	30.8	74821	9	AC109642	AC109642 Homo sapi	C	190	16	30.8	226874	2	AC099458	AC099458 Homo sapi
C	118	16	30.8	78219	3	AC004422	AC004422 Drosophi	C	191	16	30.8	229560	2	BX005404	BX005404 Danio rer
C	119	16	30.8	77137	8	AC027135	AC027135 Arabidops	C	192	16	30.8	234133	2	AC095750	AC095750 Rattus no
C	120	16	30.8	80139	8	AC020009	AC020009 Drosophi	C	193	16	30.8	239781	2	AC106535	AC106535 Rattus no
C	121	16	30.8	83876	2	AB010075	AB010075 Arabidops	C	194	16	30.8	245210	2	AC137771	AC137771 Homo sapi
C	122	16	30.8	84440	8	AP005816	AP005816 Oryza sat	C	195	16	30.8	246279	2	BX470224	BX470224 Danio rer
C	123	16	30.8	92000	2	AC105319	AC105319 Oryza sat	C	196	16	30.8	248366	2	AC113164	AC113164 Rattus no
C	124	16	30.8	96035	2	AC091510	AC091510 Leishmani	C	197	16	30.8	248366	2	AC096062	AC096062 Rattus no
C	125	16	30.8	104170	2	AC091510	AC091510 Leishmani	C	198	16	30.8	258950	2	AC099160	AC099160 Rattus no
C	126	16	30.8	110000	2	AC113913	AC113913 Drosophi	C	199	16	30.8	264629	2	AC095776	AC095776 Rattus no
C	127	16	30.8	110000	2	LMFLCHR12_13	Continuation (14 o	C	200	16	30.8	266881	1	AE017025	AE017025 Bacillus
C	128	16	30.8	110000	2	LMFLCHR12_14	Continuation (15 o	C	201	16	30.8	290714	1	AE017025	AE017025 Bacillus
C	129	16	30.8	110000	2	LMFLCHR31_20	Continuation (21 o	C	202	16	30.8	295500	1	AP005954	AP005954 Bradyrhiz
C	130	16	30.8	110000	2	LMFLCHR36_03	Continuation (4 of	C	203	16	30.8	295775	2	AC096061	AC096061 Oryza sat
C	131	16	30.8	118086	8	ATF6H11	AL021684 Arabidops	C	204	16	30.8	297470	2	AC125817	AC125817 Rattus no
C	132	16	30.8	127545	2	BX324148	BX324148 Danio rer	C	205	16	30.8	299850	2	AP005963	AP005963 Bradyrhiz
C	133	16	30.8	129430	9	AC008905	AC008905 Homo sapi	C	206	16	30.8	300029	8	AE017103	AE017103 Oryza sat
C	134	16	30.8	133253	2	AC131246	AC131246 Oryza sat	C	207	16	30.8	301250	3	AE003500	AE003500 Drosophi
C	135	16	30.8	135609	8	AC079021	AC079021 Oryza sat	C	208	16	30.8	303435	1	AE003500	AE003500 Drosophi
C	136	16	30.8	135876	8	AC113337	AC113337 Genomica s	C	209	16	30.8	306067	8	RME603645	RME603645 Rhizobium
C	137	16	30.8	138335	8	AC119148	AC119148 Oryza sat	C	210	16	30.8	306067	8	AE017063	AE017063 Oryza sat
C	138	16	30.8	139072	9	AP000485	AP000485 Homo sapi	C	211	16	30.8	307053	8	AE017079	AE017079 Oryza sat

C 212	16	30.8	314957	3	AE003581	Drosophila
C 213	16	30.8	346294	1	AP002999	Mesorhizo
C 214	16	30.8	349498	1	AP003002	Mesorhizo
C 215	16	30.8	349980	6	AX493782	Sequence
C 216	16	30.8	349980	6	AX553949	Sequence
C 217	16	28.8	226	8	AY020310	Oryza sat
C 218	15	28.8	260	9	HS71E3R	Z65680 H. sapiens C
C 219	15	28.8	272	14	BLVNADMA	119257 Bovine Leuk
C 220	15	28.8	329	14	ASSVCG	X17696 Apple scar
C 221	15	28.8	330	3	AY165174	AY165174 Tomcus p
C 222	15	28.8	330	14	ASSVXK	Y00435 Apple scar
C 223	15	28.8	331	14	AF421195	AF421195 Apple sca
C 224	15	28.8	331	14	DAVDCG	X71599 Dapple Appl
C 225	15	28.8	423	16	AR233241	AR233241 Sequence
C 226	15	28.8	444	1	UBA306168	AJ306168 unculture
C 227	15	28.8	448	6	AR128198	AR128198 Sequence
C 228	15	28.8	471	6	BD025746	BD025746 Sequence
C 229	15	28.8	535	11	G83520	G83520 S208P674FH
C 230	15	28.8	536	12	AY199102	AY199102 Arabidops
C 231	15	28.8	571	8	ZM2EM2	X85335 Z. mays mRNA
C 232	15	28.8	571	11	BV016835	BV016835 S208P6272
C 233	15	28.8	667	6	BD147972	BD147972 Primer fo
C 234	15	28.8	671	9	HSA331133	AJ331133 Homo sapi
C 235	15	28.8	680	8	MZEMADSC	L46399 Zea mays MA
C 236	15	28.8	694	3	AF438233	AF438233 Leishmani
C 237	15	28.8	753	6	BD150581	BD150581 Primer fo
C 238	15	28.8	859	4	RABKCHVG	L01789 Oryzicolaus
C 239	15	28.8	866	8	ZM2EM1	X85334 Z. mays mRNA
C 240	15	28.8	903	8	ZM2EM3	X85336 Z. mays mRNA
C 241	15	28.8	927	9	BC017347	BC017347 Homo sapi
C 242	15	28.8	1032	6	AX653040	AX653040 Sequence
C 243	15	28.8	1053	1	UBA494772	AJ944772 Unculture
C 244	15	28.8	1054	10	BC021770	BC021770 Mus muscu
C 245	15	28.8	1113	3	LME313202	AJ133022 Leishmani
C 246	15	28.8	1115	5	CHKMP4	AJ133022 Leishmani
C 247	15	28.8	1190	6	AX654314	D10599 Chicken WRF
C 248	15	28.8	1201	8	AF264731	AF264731 Oryza sat
C 249	15	28.8	1314	8	AF327528	AF327528 Arabidops
C 250	15	28.8	1335	8	AF324719	AF324719 Arabidops
C 251	15	28.8	1351	5	AF232670	AF232670 Protocpter
C 252	15	28.8	1374	1	HS222E131	HL45014 Novel hum
C 253	15	28.8	1491	1	SFU09533	U90533 Streptomyc
C 254	15	28.8	1561	8	AY056192	AY056192 Arabidops
C 255	15	28.8	1597	8	AF207593	AF207593 Arabidops
C 256	15	28.8	1724	5	CJUI2641	UI2641 Cotunux ja
C 257	15	28.8	1769	9	BC004290	BC004290 Homo sapi
C 258	15	28.8	1777	3	AK112664	AK112664 Ciona int
C 259	15	28.8	2000	6	AX508283	AX508283 Sequence
C 260	15	28.8	2024	10	BC027051	BC027051 Mus muscu
C 261	15	28.8	2034	8	AB080193	AB080193 Pium sat
C 262	15	28.8	2078	9	BC033099	BC033099 Homo sapi
C 263	15	28.8	2112	9	BC041802	BC041802 Homo sapi
C 264	15	28.8	2184	8	BC030005	BC030005 Homo sapi
C 265	15	28.8	2212	8	AY140043	AY140043 Arabidops
C 266	15	28.8	2336	3	AY071759	AY071759 Drosophila
C 267	15	28.8	2467	5	PCU44726	U44726 Penicillium
C 268	15	28.8	2494	5	AF011356	AF011356 Gallus ga
C 269	15	28.8	2520	9	AF206328	AF206328 Homo sapi
C 270	15	28.8	2554	1	BBR303058	AJ303058 Bordetell
C 271	15	28.8	2554	1	BBR303059	AJ303059 Bordetell
C 272	15	28.8	2554	1	BBR303061	AJ303061 Bordetell
C 273	15	28.8	2554	1	BBR303061	AJ303061 Bordetell
C 274	15	28.8	2554	1	BBR303062	AJ303062 Bordetell
C 275	15	28.8	2554	1	BBR303063	AJ303063 Bordetell
C 276	15	28.8	2554	1	BBR303064	AJ303064 Bordetell
C 277	15	28.8	2554	1	BBR303066	AJ303066 Bordetell
C 278	15	28.8	2554	1	BBR303071	AJ303071 Bordetell
C 279	15	28.8	2556	6	BD159688	BD159688 Primer fo
C 280	15	28.8	2556	6	AK022879	AK022879 Homo sapi
C 281	15	28.8	2580	6	BC008766	BC008766 Homo sapi
C 282	15	28.8	2598	6	AX714945	AX714945 Sequence
C 283	15	28.8	2598	9	AK057694	AK057694 Homo sapi
C 284	15	28.8	2617	6	BD158744	BD158744 Primer fo
C 285	15	28.8	2617	9	AK023976	AK023976 Homo sapi
C 286	15	28.8	2650	3	AF144322	AF144322 Giardia i
C 287	15	28.8	2693	9	BC014906	BC014906 Homo sapi
C 288	15	28.8	2697	9	BC044887	BC044887 Homo sapi
C 289	15	28.8	2702	9	AF203474	AF203474 Homo sapi
C 290	15	28.8	2790	6	AX333036	AX333036 Sequence
C 291	15	28.8	2790	6	AX335818	AX335818 Sequence
C 292	15	28.8	2790	6	HSU10868	U10868 Human aldeh
C 293	15	28.8	2794	8	BT004338	BT004338 Arabidops
C 294	15	28.8	2808	9	BC001655	BC001655 Homo sapi
C 295	15	28.8	2808	9	BC001843	BC001843 Homo sapi
C 296	15	28.8	2821	8	AF443119	AF443119 Ustilago
C 297	15	28.8	2837	9	BC013584	BC013584 Homo sapi
C 298	15	28.8	2947	1	AF239164	AF239164 Rhizobium
C 299	15	28.8	2964	9	AK058012	AK058012 Homo sapi
C 300	15	28.8	2969	8	AY059771	AY059771 Arabidops
C 301	15	28.8	3391	1	TFALAS	X95571 T. ferrooxid
C 302	15	28.8	3514	3	DMSING3	X17550 D. melanoga
C 303	15	28.8	3866	3	BT003619	BT003619 Drosophila
C 304	15	28.8	3963	6	AX260206	AX260206 Sequence
C 305	15	28.8	4038	1	MXALON	D12923 Myxococcus
C 306	15	28.8	4198	6	AX067966	AX067966 Sequence
C 307	15	28.8	4198	8	ATENRA	Y13860 Arabidopsis
C 308	15	28.8	4359	3	AY094956	AY094956 Drosophila
C 309	15	28.8	4463	8	AF355592	AF355592 Filobaasid
C 310	15	28.8	4475	8	ZM2EMAGEN	X91882 Z. mays ZEM
C 311	15	28.8	4516	3	AB037841	AB037841 Homo sapi
C 312	15	28.8	4750	3	AF197953	AF197953 Topoplaem
C 313	15	28.8	4800	1	AB044351	AB044351 Pseudomon
C 314	15	28.8	4842	3	AF132179	AF132179 Drosophila
C 315	15	28.8	4945	10	RNO550165	AJ550165 Rattus no
C 316	15	28.8	4986	2	AC019465	AC019465 Drosophila
C 317	15	28.8	5118	6	AF79572	AF79572 Sequence 3
C 318	15	28.8	5118	6	AF175232	AF175232 Sequence
C 319	15	28.8	5170	10	RNO426052	AJ426052 Rattus no
C 320	15	28.8	5334	1	BBCYAG	ZJ7112 Bordetella
C 321	15	28.8	5369	1	BPA249835	AJ249835 Bordetella
C 322	15	28.8	5419	3	DME238847	AJ38847 Drosophila
C 323	15	28.8	5694	14	HSC2CHOM	UJ2062 Cercopithec
C 324	15	28.8	5721	6	A25910	A25910 B. bronchis
C 325	15	28.8	5962	1	MBS10OPER	Y13228 Mycobacteri
C 326	15	28.8	6000	6	AF433876	AF433876 Giardia i
C 327	15	28.8	6059	6	AX281301	AX281301 Sequence
C 328	15	28.8	6078	3	AB061681	AB061681 Drosophila
C 329	15	28.8	6441	6	A14850	A14850 B. pertussis
C 330	15	28.8	6441	6	AF9570	AF9570 Sequence 1
C 331	15	28.8	6441	6	AR175231	AR175231 Sequence
C 332	15	28.8	6441	6	BPCYA	Y00545 Bordetella
C 333	15	28.8	6443	6	A07292	A07292 B. pertussis
C 334	15	28.8	6443	6	A32503	A32503 B. anthracis
C 335	15	28.8	6991	8	AB084921	AB084921 Paecilomy
C 336	15	28.8	7646	1	AF116906	AF116906 Rhodococ
C 337	15	28.8	9349	1	AF173961	AF173961 Thauera a
C 338	15	28.8	9405	6	AX399934	AX399934 Sequence
C 339	15	28.8	9490	4	AF091049	AF091049 Canis fam
C 340	15	28.8	9926	9	AF257305	AF257305 Homo sapi
C 341	15	28.8	9959	1	AE014682	AE014682 Bifidobac
C 342	15	28.8	10057	1	AE004871	AE004871 Pseudomon
C 343	15	28.8	10151	8	AF484941	AF484941 Fusarium
C 344	15	28.8	10205	1	AE004830	AE004830 Pseudomon
C 345	15	28.8	10210	1	AE014776	AE014776 Bifidobac
C 346	15	28.8	10317	1	AE011691	AE011691 Xanthomon
C 347	15	28.8	10452	1	AE012453	AE012453 Xanthomon
C 348	15	28.8	10510	1	AE012870	AE012870 Chlorobiu
C 349	15	28.8	10510	1	AE000088	AE000088 Rhizobium
C 350	15	28.8	10734	1	AE012845	AE012845 Myxococcu
C 351	15	28.8	10847	1	AF127082	AF127082 Chlorobiu
C 352	15	28.8	10930	1	AE012926	AE012926 Chlorobiu
C 353	15	28.8	11415	1	AE004889	AE004889 Pseudomon
C 354	15	28.8	11431	1	AF372655	AF372655 Rhizobium
C 355	15	28.8	11455	1	AE001906	AE001906 Deinococc
C 356	15	28.8	11462	1	AE011895	AE011895 Xanthomon
C 357	15	28.8	11684	1	AE005903	AE005903 Caulobact

C 358	15	28.8	11847	6	AX254620	Sequence	C 431	15	28.8	100986	8	AC105363	Oryza sat
C 359	15	28.8	11855	2	AC014860	AE014860 Drosophila	C 432	15	28.8	103874	2	AC123912	Human sapi
C 360	15	28.8	12201	1	AE012255	Xanthomon	C 433	15	28.8	103814	2	AC009210	AE009210 Drosophila
C 361	15	28.8	12277	1	AEF34016	Streptomy	C 434	15	28.8	103814	2	AC009210	Human sapi
C 362	15	28.8	12736	1	AE014798	Bifidobac	C 435	15	28.8	107661	9	AL162734	AL162734 Human DNA
C 363	15	28.8	12787	1	AE0011720	AE0011720 Thermotog	C 436	15	28.8	107947	8	NCEB7N14	AL669986 Neurospor
C 364	15	28.8	12789	1	AE012214	Xanthomon	C 437	15	28.8	108582	8	AP006098	AP006098 Lotus japa
C 365	15	28.8	14435	2	AC013324	AC013324 Drosophila	C 438	15	28.8	110000	2	AC129386	Continuati
C 366	15	28.8	14828	1	AE004676	AE004676 Pseudomon	C 439	15	28.8	110000	2	LMF1CHR16_0	Continuati
C 367	15	28.8	15417	1	AE006866	Mycobacte	C 440	15	28.8	110000	2	LMF1CHR16_01	Continuati
C 368	15	28.8	15573	1	AE006866	Mycobacte	C 441	15	28.8	110000	2	LMF1CHR22_07	Continuati
C 369	15	28.8	16124	6	AX024277	Sequence	C 442	15	28.8	110000	2	LMF1CHR33_13	Continuati
C 370	15	28.8	16623	1	AE005661	Escherich	C 443	15	28.8	110000	2	LMF1CHR33_23	Continuati
C 371	15	28.8	16623	1	AE007074	Mycobacte	C 444	15	28.8	110000	2	LMF1CHR33_24	Continuati
C 372	15	28.8	16929	1	AE006917	Mycobacte	C 445	15	28.8	110000	2	LMF1CHR33_13	Continuati
C 373	15	28.8	18282	1	AF141932	AF141932 Rhizobium	C 446	15	28.8	110000	2	LMF1CHR33_15	Continuati
C 374	15	28.8	23996	2	AC012981	AC012981 Drosophila	C 447	15	28.8	110000	2	LMF1CHR33_15	Continuati
C 375	15	28.8	24655	1	AL162498	Human DNA	C 448	15	28.8	110000	8	AC145127_14	Continuati
C 376	15	28.8	25600	3	U80442	U80442 Caenorhabin	C 449	15	28.8	110985	8	AP004578	AP004578
C 377	15	28.8	29359	8	AC005931	AC005931 Leishmani	C 450	15	28.8	113459	2	AP006592	AP006592
C 378	15	28.8	29538	8	APU85909	U85909 Aureobasidi	C 451	15	28.8	114000	8	AC125412	AC125412
C 379	15	28.8	30098	2	AC079026	AC079026 Leishmani	C 452	15	28.8	115446	2	AC120402	AC120402
C 380	15	28.8	31638	1	RMEXP8NS	Z79692 R. meilloti	C 453	15	28.8	120145	2	AC098842	AC098842
C 381	15	28.8	31880	6	AR204546	AR204546 Sequence	C 454	15	28.8	121038	8	AP003342	AP003342
C 382	15	28.8	32765	9	HS363412	Human DNA	C 455	15	28.8	121160	8	AC005970	AC005970
C 383	15	28.8	35540	1	AF389112	AF389112 Rhizobium	C 456	15	28.8	121163	9	AC106879	AC106879
C 384	15	28.8	36526	1	MSCGY42	AD000005 Mycobacteri	C 457	15	28.8	122116	8	AC137634	AC137634
C 385	15	28.8	36804	1	MTCY210	Z84395 Mycobacteri	C 458	15	28.8	122390	2	AC135918	AC135918
C 386	15	28.8	37991	9	AP000565	AP000565 Homo sapi	C 459	15	28.8	123360	2	HS222813	HS222813
C 387	15	28.8	38110	1	MTCY427	Z70692 Mycobacteri	C 460	15	28.8	124205	8	CNS08CCQ	CNS08CCQ
C 388	15	28.8	38653	6	AR285078	AR285078 Sequence	C 461	15	28.8	124290	2	AP006059	AP006059 Oryza sat
C 389	15	28.8	38721	1	MSGY219	AD000013 Mycobacte	C 462	15	28.8	124711	8	AP003808	AP003808
C 390	15	28.8	38874	3	LMFLJ377	AL138558 Leishmani	C 463	15	28.8	126731	8	AP003379	AP003379
C 391	15	28.8	39160	1	MTCY10H4	Z80213 Mycobacteri	C 464	15	28.8	126657	8	AC006234	AC006234 Arabidops
C 392	15	28.8	40013	3	LMFLJ508	AL596287 Leishmani	C 465	15	28.8	126667	8	AC006234	AC006234 Arabidops
C 393	15	28.8	40374	3	LMFLJ302	AL359721 Leishmani	C 466	15	28.8	131768	2	AP005476	AP005476
C 394	15	28.8	41097	1	AF016585	AF016585 Streptomy	C 467	15	28.8	132575	10	AC117261	AC117261 Mus musc
C 395	15	28.8	41285	3	AC005930	AC005930 Leishmani	C 468	15	28.8	132790	9	HS167A14	AC117261 Human DNA s
C 396	15	28.8	41400	8	AP005254	AP005254 Oryza sat	C 469	15	28.8	133790	8	AP003842	AP003842 Oryza sat
C 397	15	28.8	42498	1	MLCL622	Z95398 Mycobacteri	C 470	15	28.8	134367	8	AP004888	AP004888 Oryza sat
C 398	15	28.8	43417	2	BX294390	BX294390 Homo sapi	C 471	15	28.8	135782	2	AC125385	AC125385 Oryza sat
C 399	15	28.8	44366	1	AP003016	AP003016 Mesophilu	C 472	15	28.8	138084	2	AC099074	AC099074 Oryza sat
C 400	15	28.8	54735	5	AF456419	AF456419 Takiflugu	C 473	15	28.8	139482	2	AC113333	AC113333 Oryza sat
C 401	15	28.8	55933	1	BSMJ2571	AJ002571 Bacilli	C 474	15	28.8	139520	5	AL935199	AL935199 Zebrafish
C 402	15	28.8	59414	1	AC091029	AC091029 Homo sapi	C 475	15	28.8	139645	2	AC105932	AC105932 Oryza sat
C 403	15	28.8	59580	9	AC106875	AC106875 Homo sapi	C 476	15	28.8	139841	2	OSJN00054	AL066621 Oryza sat
C 404	15	28.8	63809	2	AC135769	AC135769 Rattus no	C 477	15	28.8	140583	2	AP005413	AP005413 Oryza sat
C 405	15	28.8	66800	10	AL732607	AL732607 Mouse DNA	C 478	15	28.8	140688	2	AP003801	AP003801 Oryza sat
C 406	15	28.8	68750	1	AF210843	AF210843 Sorangium	C 479	15	28.8	142133	2	AC135014	AC135014
C 407	15	28.8	68750	6	AR193029	AR193029 Sequence	C 480	15	28.8	142184	2	AC037428	AC037428 Homo sapi
C 408	15	28.8	68750	6	AR199551	AR199551 Sequence	C 481	15	28.8	142195	2	AC010669	AC010669 Drosophila
C 409	15	28.8	68750	6	AR199559	AR199559 Sequence	C 482	15	28.8	143117	2	AP005872	AP005872 Oryza sat
C 410	15	28.8	68750	6	AR199567	AR199567 Sequence	C 483	15	28.8	143117	2	AP005872	AP005872 Oryza sat
C 411	15	28.8	68750	6	AR201097	AR201097 Sequence	C 484	15	28.8	143504	9	AC093610	AC093610 Homo sapi
C 412	15	28.8	68750	6	AR208671	AR208671 Sequence	C 485	15	28.8	143668	8	AP004268	AP004268 Oryza sat
C 413	15	28.8	71989	5	AR172664	AR172664 Sequence	C 486	15	28.8	143668	8	AC132485	AC132485 Oryza sat
C 414	15	28.8	72887	2	AC017882	AC017882 Drosophila	C 487	15	28.8	143809	2	AC130604	AC130604 Oryza sat
C 415	15	28.8	79422	2	AC022750	AC022750 Homo sapi	C 488	15	28.8	145003	2	AL935049	AL935049 Danio rer
C 416	15	28.8	80609	1	AF116907	AF116907 Rhodococc	C 489	15	28.8	145339	2	AP005608	AP005608 Oryza sat
C 417	15	28.8	80610	1	AP001204	AP001204 Rhodococc	C 490	15	28.8	145414	9	HS78F24	AL022336 Human DNA
C 418	15	28.8	81777	2	AC144927	AC144927 Mus muscu	C 491	15	28.8	145662	8	CNS08CAV	AL054498 Oryza sat
C 419	15	28.8	82010	8	AB023042	AB023042 Arabidops	C 492	15	28.8	146568	8	AC105364	AC105364 Oryza sat
C 420	15	28.8	83925	8	NCEB23121	NCEB23121 Neurospor	C 493	15	28.8	146576	8	AC074282	AC074282 Oryza sat
C 421	15	28.8	84462	8	AB023038	AB023038 Arabidops	C 494	15	28.8	146936	2	AP005010	AP005010 Oryza sat
C 422	15	28.8	84518	3	AC135132	Continuati	C 495	15	28.8	147348	2	AC132489	AC132489 Oryza sat
C 423	15	28.8	85167	3	AC131532	Continuati	C 496	15	28.8	147750	2	AC018761	AC018761 Homo sapi
C 424	15	28.8	88277	2	AC015811	AC015811 Homo sapi	C 497	15	28.8	149098	8	AP003296	AP003296 Oryza sat
C 425	15	28.8	92053	2	AC012720	AC012720 Drosophila	C 498	15	28.8	151798	2	AC0044819	AC0044819 Homo sapi
C 426	15	28.8	93475	5	AL672088	AL672088 zebrafish	C 499	15	28.8	152075	2	CNS07EFL	AL034006 Oryza sat
C 427	15	28.8	94851	9	AC004923	AC004923 Homo sapi	C 500	15	28.8	152081	2	AC139631	AC139631 Bos tauru
C 428	15	28.8	95521	2	AC134499	AC134499 Rattus no	C 501	15	28.8	152081	2	CNS08C82	AL732376 Oryza sat
C 429	15	28.8	96310	2	AL136444	AL136444 Human DNA	C 502	15	28.8	152274	8	AP004669	AP004669 Oryza sat
C 430	15	28.8	96468	2	AC006902	AC006902 Caenorhabin	C 503	15	28.8	152423	8	AC079853	AC079853 Oryza sat

C 504	15	28.8 152736	8	AP003566	577	15	28.8 189862	2	AC108317	AC108317 Rattus no
C 505	15	28.8 152898	2	AC134432	C 578	15	28.8 193444	8	AF229199	AF229199 Oryza sat
C 506	15	28.8 152945	2	AC130606	C 579	15	28.8 194361	8	AC092389	AC092389 Oryza sat
C 507	15	28.8 153464	2	AC123391	C 580	15	28.8 194371	2	AC135453	AC135453 Rattus no
C 508	15	28.8 153802	9	AL589949	C 581	15	28.8 195859	14	AF281817	AF281817 Tupiaia he
C 509	15	28.8 153926	9	AL589949	C 582	15	28.8 196050	2	AL646058	AL646058 Ralstonia
C 510	15	28.8 154181	8	AC131667	C 583	15	28.8 196252	2	BX324145	BX324145 Dantio rer
C 511	15	28.8 154862	2	AC139310	C 584	15	28.8 198456	2	AC090443	AC090443 Mus muscu
C 512	15	28.8 155172	2	AC021110	C 585	15	28.8 199745	5	AL805905	AL805905 Zebrafish
C 513	15	28.8 157920	2	AP003564	C 586	15	28.8 200316	2	BX004990	BX004990 Dantio rer
C 514	15	28.8 158146	9	AP003564	C 587	15	28.8 200628	10	AL670720	AL670720 Mouse DNA
C 515	15	28.8 158508	2	AC063334	C 588	15	28.8 200868	3	AC010030	AC010030 Drosophill
C 516	15	28.8 158826	2	AC140156	C 589	15	28.8 202080	2	AC113906	AC113906 Rattus no
C 517	15	28.8 158839	8	AC090882	C 590	15	28.8 202080	2	AC129955	AC129955 Bos tauru
C 518	15	28.8 159391	2	AC069199	C 591	15	28.8 203050	1	AL646078	AL646078 Ralstonia
C 519	15	28.8 160931	3	AC093498	C 592	15	28.8 204010	2	AC139412	AC139412 Rattus no
C 520	15	28.8 160984	2	AC021281	C 593	15	28.8 204050	9	AC034259	AC034259 Ralstonia
C 521	15	28.8 161601	3	AC007594	C 594	15	28.8 206192	2	AC018039	AC018039 Homo sapi
C 522	15	28.8 161720	8	AP003281	C 595	15	28.8 207684	2	AC018039	AC018039 Drosophill
C 523	15	28.8 162719	2	AC021221	C 596	15	28.8 208230	1	BSUB0008	BSUB0008 Homo sapi
C 524	15	28.8 162721	10	AC079638	C 597	15	28.8 212840	2	AC135223	AC135223 Gallus ga
C 525	15	28.8 163012	2	AC142104	C 598	15	28.8 213984	2	AC135111	AC135111 Mus muscu
C 526	15	28.8 163795	2	AC102466	C 599	15	28.8 216750	1	BSUB0007	BSUB0007 299110 Bacillus su
C 527	15	28.8 163797	2	AC102466	C 600	15	28.8 217292	9	AF288742	AF288742 Homo sapi
C 528	15	28.8 163584	5	AL953872	C 601	15	28.8 219366	2	AP001922	AP001922 Homo sapi
C 529	15	28.8 165900	2	AP003942	C 602	15	28.8 219491	2	AC092256	AC092256 Mus muscu
C 530	15	28.8 165900	2	AC011072	C 603	15	28.8 223099	2	AC107506	AC107506 Rattus no
C 531	15	28.8 166160	3	BX005141	C 604	15	28.8 223326	2	AC127905	AC127905 Rattus no
C 532	15	28.8 166710	2	AC005141	C 605	15	28.8 223438	2	AC117094	AC117094 Rattus no
C 533	15	28.8 166862	9	CNS01DTS	C 606	15	28.8 224786	2	AC118501	AC118501 Rattus no
C 534	15	28.8 166946	2	AC116781	C 607	15	28.8 225041	2	AC120658	AC120658 Rattus no
C 535	15	28.8 167239	3	AC007668	C 608	15	28.8 225684	10	AL845492	AL845492 Mouse DNA
C 536	15	28.8 167658	2	AC123942	C 609	15	28.8 225730	2	AC119454	AC119454 Rattus no
C 537	15	28.8 168259	2	AC067842	C 610	15	28.8 227459	2	AC127039	AC127039 Rattus no
C 538	15	28.8 168764	8	AP002569	C 611	15	28.8 227901	2	AC141209	AC141209 Mus muscu
C 539	15	28.8 168764	8	AC087726	C 612	15	28.8 228350	2	AC141058	AC141058 Homo sapi
C 540	15	28.8 169137	2	AC117741	C 613	15	28.8 229889	2	AC136274	AC136274 Rattus no
C 541	15	28.8 169142	8	AC053514	C 614	15	28.8 230237	2	AC131463	AC131463 Rattus no
C 542	15	28.8 169258	9	AC022013	C 615	15	28.8 231168	2	AC140268	AC140268 Mus muscu
C 543	15	28.8 169728	8	AC091787	C 616	15	28.8 231741	3	AE003705	AE003705 Drosophill
C 544	15	28.8 169951	2	AC015654	C 617	15	28.8 232993	2	AC098761	AC098761 Rattus no
C 545	15	28.8 171861	2	AC104824	C 618	15	28.8 233355	2	AC095215	AC095215 Rattus no
C 546	15	28.8 172190	3	AC010668	C 619	15	28.8 234580	2	AC130855	AC130855 Rattus no
C 547	15	28.8 172647	8	AP004258	C 620	15	28.8 235188	2	AC108344	AC108344 Rattus no
C 548	15	28.8 173147	2	BX294117	C 621	15	28.8 235021	2	AC094509	AC094509 Rattus no
C 549	15	28.8 173728	9	AC018507	C 622	15	28.8 236062	2	AC105688	AC105688 Rattus no
C 550	15	28.8 175192	3	AC015076	C 623	15	28.8 236165	6	A93003	A93003 Sequence 2
C 551	15	28.8 176295	2	AC026176	C 624	15	28.8 237119	2	AC095707	AC095707 Drosophill
C 552	15	28.8 177189	2	AC142049	C 625	15	28.8 237426	2	AC095707	AC095707 Rattus no
C 553	15	28.8 178009	2	AC119549	C 626	15	28.8 237426	2	AC094219	AC094219 Rattus no
C 554	15	28.8 178073	1	SC059046	C 627	15	28.8 237426	2	AC106678	AC106678 Rattus no
C 555	15	28.8 178249	2	AP006058	C 628	15	28.8 237426	2	AC129456	AC129456 Rattus no
C 556	15	28.8 178473	10	AL672146	C 629	15	28.8 240561	2	AC129456	AC129456 Rattus no
C 557	15	28.8 178776	2	AC068217	C 630	15	28.8 241432	10	AL589661	AL589661 Chimpanze
C 558	15	28.8 179064	2	AL353807	C 631	15	28.8 241432	2	AC136816	AC136816 Rattus no
C 559	15	28.8 179439	9	AC026462	C 632	15	28.8 241692	2	AC136816	AC136816 Rattus no
C 560	15	28.8 179640	2	AC026462	C 633	15	28.8 243114	10	AL672244	AL672244 Mouse DNA
C 561	15	28.8 181178	3	AC023713	C 634	15	28.8 244625	5	AL928716	AL928716 Zebrafish
C 562	15	28.8 181528	3	AC010562	C 635	15	28.8 246647	2	AC130098	AC130098 Rattus no
C 563	15	28.8 183048	2	AC027439	C 636	15	28.8 247296	2	AC127145	AC127145 Rattus no
C 564	15	28.8 183094	10	AC087903	C 637	15	28.8 247296	2	AC111661	AC111661 Rattus no
C 565	15	28.8 183439	3	AC007809	C 638	15	28.8 253505	1	AP000007	AP000007 Pyrococcus
C 566	15	28.8 184600	2	AC117347	C 639	15	28.8 253505	10	AL589699	AL589699 Mouse DNA
C 567	15	28.8 184940	2	AL672264	C 640	15	28.8 253505	2	AC095671	AC095671 Rattus no
C 568	15	28.8 186487	8	BX248402	C 641	15	28.8 259606	2	AC099416	AC099416 Mus muscu
C 569	15	28.8 186609	8	MP0MTG	C 642	15	28.8 259606	2	AC109119	AC109119 Rattus no
C 570	15	28.8 187332	9	AC101292	C 643	15	28.8 260050	1	AC094000	AC094000 Rattus no
C 571	15	28.8 187386	2	AC099507	C 644	15	28.8 260050	2	AC125564	AC125564 Rattus no
C 572	15	28.8 187439	2	AC107710	C 645	15	28.8 263357	2	AC125564	AC125564 Rattus no
C 573	15	28.8 187516	2	AL928929	C 646	15	28.8 265467	3	AE003696	AE003696 Drosophill
C 574	15	28.8 189043	8	AP004366	C 647	15	28.8 265467	2	AC127791	AC127791 Rattus no
C 575	15	28.8 189050	1	AL646066	C 648	15	28.8 265467	2	AC127791	AC127791 Rattus no
C 576	15	28.8 189219	2	AC009680	C 649	15	28.8 266035	2	AC127791	AC127791 Rattus no

C 650	15	28.8	269203	1	MLEPRTN10	AL581926	Mycobacte	C 723	14	26.9	324	1	AF016612	AF016612	Unidentif
C 651	15	28.8	270061	2	AC094481	AC094481	Rattus no	C 724	14	26.9	345	6	BD031435	BD031435	Sequence
C 652	15	28.8	274059	2	AC123367	AC123367	Rattus no	C 725	14	26.9	347	6	DME30074	DME30074	Drosophil
C 653	15	28.8	275690	2	AC096659	AC096659	Rattus no	C 726	14	26.9	357	8	AY198954	AY198954	Arabidops
C 654	15	28.8	277615	2	AC125485	AC125485	Mus muscu	C 727	14	26.9	358	8	AF071124	AF071124	Zea mays
C 655	15	28.8	281339	2	AE003547	AE003547	Drosophil	C 728	14	26.9	373	11	AU026831	AU026831	Rattus no
C 656	15	28.8	285494	3	AE003560	AE003560	Drosophil	C 729	14	26.9	390	6	AX069894	AX069894	Sequence
C 657	15	28.8	295150	1	SC0939126	SC0939126	Streptomy	C 730	14	26.9	414	1	AF414652	AF414652	Unculture
C 658	15	28.8	295500	1	AP005954	AP005954	Bradyrhiz	C 731	14	26.9	423	3	AF414649	AF414649	Unculture
C 659	15	28.8	298300	1	AP005025	AP005025	Streptomy	C 732	14	26.9	427	3	DME271142	DME271142	Drosophil
C 660	15	28.8	298550	1	AP005961	AP005961	Bradyrhiz	C 733	14	26.9	427	6	AX363569	AX363569	Sequence
C 661	15	28.8	299175	1	AP005023	AP005023	Streptomy	C 734	14	26.9	444	3	AF210725	AF210725	Anopheles
C 662	15	28.8	299600	1	AP005942	AP005942	Bradyrhiz	C 735	14	26.9	466	3	DME300070	DME300070	Drosophil
C 663	15	28.8	299925	1	AP005039	AP005039	Streptomy	C 736	14	26.9	510	3	DME300075	DME300075	Drosophil
C 664	15	28.8	300029	8	AE017051	AE017051	Oryza sat	C 737	14	26.9	512	8	AY183067	AY183067	Triticolum
C 665	15	28.8	300029	8	AE017067	AE017067	Oryza sat	C 738	14	26.9	515	6	AX472762	AX472762	Sequence
C 666	15	28.8	300029	8	AE017115	AE017115	Oryza sat	C 739	14	26.9	533	6	AX660612	AX660612	Sequence
C 667	15	28.8	300200	1	AP005962	AP005962	Bradyrhiz	C 740	14	26.9	537	6	AX094768	AX094768	Sequence
C 668	15	28.8	300242	1	AE016790	AE016790	Pseudomon	C 741	14	26.9	543	8	AY083989	AY083989	Zea mays
C 669	15	28.8	300550	1	AP005021	AP005021	Streptomy	C 742	14	26.9	567	3	DME300066	DME300066	Drosophil
C 670	15	28.8	300600	1	AP005936	AP005936	Bradyrhiz	C 743	14	26.9	568	3	DME300072	DME300072	Drosophil
C 671	15	28.8	300695	2	AC079431	AC079431	Mus muscu	C 744	14	26.9	581	8	AF424659	AF424659	Phytocith
C 672	15	28.8	302675	1	AP005024	AP005024	Streptomy	C 745	14	26.9	581	8	AF424687	AF424687	Phytocith
C 673	15	28.8	302914	2	AC136117	AC136117	Rattus no	C 746	14	26.9	589	8	AY183070	AY183070	Triticolum
C 674	15	28.8	303335	3	AE003538	AE003538	Drosophil	C 747	14	26.9	589	11	G98148	G98148	Triticolum
C 675	15	28.8	303450	1	SC0939130	SC0939130	Streptomy	C 748	14	26.9	592	8	AY182999	AY182999	Triticolum
C 676	15	28.8	306050	1	BX248341	BX248341	Mycobacte	C 749	14	26.9	593	8	AY178301	AY178301	Triticolum
C 677	15	28.8	306050	1	RME603645	RME603645	Rhizobium	C 750	14	26.9	593	8	AY183074	AY183074	Triticolum
C 678	15	28.8	307250	1	AP005219	AP005219	Corynebac	C 751	14	26.9	594	8	AY183076	AY183076	Triticolum
C 679	15	28.8	310021	2	AC092883	AC092883	Homo sapi	C 752	14	26.9	597	10	BC051997	BC051997	Mus muscu
C 680	15	28.8	310050	1	RME603642	RME603642	Rhizobium	C 753	14	26.9	599	11	G85017	G85017	S210P6224RB
C 681	15	28.8	310325	1	AE016864	AE016864	Pseudomon	C 754	14	26.9	606	8	AY124824	AY124824	Arabidops
C 682	15	28.8	315000	1	RME603644	RME603644	Rhizobium	C 755	14	26.9	607	11	BV056693	BV056693	S212P6030
C 683	15	28.8	315362	3	AE003442	AE003442	Drosophil	C 756	14	26.9	612	6	AX412732	AX412732	Sequence
C 684	15	28.8	316961	2	AE003799	AE003799	Drosophil	C 757	14	26.9	612	6	AX505434	AX505434	Sequence
C 685	15	28.8	318505	2	AC123174	AC123174	Rattus no	C 758	14	26.9	612	6	AX589867	AX589867	Sequence
C 686	15	28.8	319749	2	AC123174	AC123174	Rattus no	C 759	14	26.9	612	8	BT008445	BT008445	Arabidops
C 687	15	28.8	320040	6	A79350	A79350	Sequence 1	C 760	14	26.9	651	10	AB008682	AB008682	Rattus no
C 688	15	28.8	320040	6	A79350	A79350	Sequence 1	C 761	14	26.9	662	8	AF339730	AF339730	Arabidops
C 689	15	28.8	320050	1	AX248336	AX248336	Mycobacte	C 762	14	26.9	691	1	PCU10242	PCU10242	Pseudomonas
C 690	15	28.8	324288	2	AC017336	AC017336	Drosophil	C 763	14	26.9	693	11	BV034435	BV034435	Sequence
C 691	15	28.8	328561	3	AE003843	AE003843	Drosophil	C 764	14	26.9	696	9	HSA334458	HSA334458	Homo sapi
C 692	15	28.8	329709	1	AP002997	AP002997	Mesothizo	C 765	14	26.9	697	3	AY113515	AY113515	Sequence
C 693	15	28.8	339193	2	AC098394	AC098394	Rattus no	C 766	14	26.9	700	10	AY158891	AY158891	Mus muscu
C 694	15	28.8	340000	9	AP001718	AP001718	Homo sapi	C 767	14	26.9	700	10	AY160973	AY160973	Mus muscu
C 695	15	28.8	340857	1	AP003010	AP003010	Mesothizo	C 768	14	26.9	700	11	BV047454	BV047454	S212P6228
C 696	15	28.8	343050	1	AX248334	AX248334	Mycobacte	C 769	14	26.9	701	10	MMU272227	MMU272227	Mus muscu
C 697	15	28.8	346510	1	AP003011	AP003011	Mesothizo	C 770	14	26.9	703	10	BC052746	BC052746	Mus muscu
C 698	15	28.8	348002	2	AC118775	AC118775	Rattus no	C 771	14	26.9	715	8	AY201517	AY201517	Arabidops
C 699	15	28.8	348077	1	AP003000	AP003000	Mesothizo	C 772	14	26.9	715	9	HSA330094	HSA330094	Homo sapi
C 700	15	28.8	348540	2	AC132503	AC132503	Rattus no	C 773	14	26.9	723	9	AB01005501	AB01005501	Homo sapi
C 701	15	28.8	349468	1	AP003002	AP003002	Mesothizo	C 774	14	26.9	729	1	RLBA11	RLBA11	Sequence
C 702	15	28.8	349619	1	AP002996	AP002996	Mesothizo	C 775	14	26.9	730	8	AY289189	AY289189	Oryza sat
C 703	15	28.8	349980	6	AX492783	AX492783	Sequence	C 776	14	26.9	734	9	HSA337241	HSA337241	Homo sapi
C 704	15	28.8	349980	6	AX492786	AX492786	Sequence	C 777	14	26.9	734	9	S46954	S46954	neurofilame
C 705	15	28.8	349980	6	AX492787	AX492787	Sequence	C 778	14	26.9	759	3	DME300065	DME300065	Drosophil
C 706	15	28.8	349980	6	AX553950	AX553950	Sequence	C 779	14	26.9	766	8	BT004428	BT004428	Arabidops
C 707	15	28.8	349980	6	AX553953	AX553953	Sequence	C 780	14	26.9	768	6	AX512898	AX512898	Sequence
C 708	15	28.8	349980	6	AX553954	AX553954	Sequence	C 781	14	26.9	779	8	AF324713	AF324713	Arabidops
C 709	14	26.9	81	3	BOU83695	U83695	Berge ovata	C 782	14	26.9	779	8	AY065086	AY065086	Arabidops
C 710	14	26.9	105	3	AF273850	AF273850	Neospora	C 783	14	26.9	797	10	RATLHR01	RATLHR01	Rat lutealni
C 711	14	26.9	115	6	BD028291	BD028291	Sequence	C 784	14	26.9	804	6	AX537934	AX537934	Sequence
C 712	14	26.9	172	9	HS67C6F	HS67C6F	H. sapiens C	C 785	14	26.9	818	6	BD148800	BD148800	Primer fo
C 713	14	26.9	190	10	AY047594	AY047594	Mertiones	C 786	14	26.9	826	11	BV026717	BV026717	S212P6555
C 714	14	26.9	195	8	AY032790	AY032790	Passiflor	C 787	14	26.9	858	11	BOV09754	BOV09754	Arabidops
C 715	14	26.9	208	4	AF085334	AF085334	Ovis arie	C 788	14	26.9	872	8	AY086954	AY086954	Arabidops
C 716	14	26.9	208	6	BD032216	BD032216	Sequence	C 789	14	26.9	875	3	AY172973	AY172973	Mastigamo
C 717	14	26.9	220	10	S40787801	S40787	lutealizing	C 790	14	26.9	889	9	HSA332150	HSA332150	Homo sapi
C 718	14	26.9	224	8	AY019862	AY019862	Oryza sat	C 791	14	26.9	897	4	AY239196	AY239196	Cyclopes
C 719	14	26.9	234	6	AR240280	AR240280	Sequence	C 792	14	26.9	909	1	EC4410311	EC4410311	Erwinia c
C 720	14	26.9	266	6	AX307534	AX307534	Sequence	C 793	14	26.9	922	10	BC012231	BC012231	Mus muscu
C 721	14	26.9	334	1	AF016610	AF016610	Unidentif	C 794	14	26.9	928	8	AF462840	AF462840	Arabidops
C 722	14	26.9	334	1	AF016611	AF016611	Unidentif	C 795	14	26.9	960	6	AR094305	AR094305	Sequence

C 796	14	26.9	968	9	HSN802426	C 869	14	26.9	1970	6	AX262402	AX262402 Sequence
C 797	14	26.9	970	5	AF027978	C 870	14	26.9	1985	8	AF424570	AF424570 Arabidops
C 798	14	26.9	984	3	GBDMP3A	C 871	14	26.9	2000	6	AX656614	AX656614 Sequence
C 799	14	26.9	991	3	AF171855	C 872	14	26.9	2008	9	HSN805524	AL834432 Homo sapi
C 800	14	26.9	1005	14	HIVU76639	C 873	14	26.9	2010	9	AK075401	AK075401 Homo sapi
C 801	14	26.9	1020	8	BT002868	C 874	14	26.9	2036	6	AX224806	AX224806 Sequence
C 802	14	26.9	1022	3	PTU66467	C 875	14	26.9	2045	6	DRONOT01	M16153 Drosophila
C 803	14	26.9	1034	6	AX720346	C 876	14	26.9	2059	6	AX714630	AX714630 Sequence
C 804	14	26.9	1052	5	AF483320	C 877	14	26.9	2059	9	AK057133	AK057133 Homo sapi
C 805	14	26.9	1057	8	AF059582	C 878	14	26.9	2063	9	BC048985	BC048985 Homo sapi
C 806	14	26.9	1058	8	AF087100	C 879	14	26.9	2098	6	AX154565	AX154565 Sequence
C 807	14	26.9	1080	6	E33745	C 880	14	26.9	2107	10	BC023820	BC023820 Mus muscu
C 808	14	26.9	1091	3	AF087104	C 881	14	26.9	2119	6	BC030591	BC030591 Homo sapi
C 809	14	26.9	1093	3	AF240865	C 882	14	26.9	2125	6	AX647819	AX647819 Sequence
C 810	14	26.9	1105	8	AF087102	C 883	14	26.9	2132	9	AF311287	AF311287 Homo sapi
C 811	14	26.9	1112	8	AF087106	C 884	14	26.9	2133	10	MM0278575	MM0278575 Mus muscu
C 812	14	26.9	1120	5	AB059834	C 885	14	26.9	2141	3	AC014879	AC014879 Drosophi1
C 813	14	26.9	1124	8	AF087107	C 886	14	26.9	2171	3	GB1302036	AJ102036 Gryllus b
C 814	14	26.9	1140	14	AY223528	C 887	14	26.9	2176	6	BD160313	BD160313 Primer fo
C 815	14	26.9	1161	6	AX451933	C 888	14	26.9	2176	9	AK024001	AK024001 Homo sapi
C 816	14	26.9	1170	3	AF268321	C 889	14	26.9	2254	3	AY058407	AY058407 Drosophi1
C 817	14	26.9	1182	6	BD016692	C 890	14	26.9	2279	3	AF190556	AF190556 Typanoso
C 818	14	26.9	1191	1	AY142858	C 891	14	26.9	2301	3	AK112563	AK112563 Ciona int
C 819	14	26.9	1228	1	AF124348	C 892	14	26.9	2306	9	SC0512793	AJ512793 Scutigera
C 820	14	26.9	1231	8	ZMA430637	C 893	14	26.9	2307	9	HSAB6470	AJ006470 Homo sapi
C 821	14	26.9	1290	8	AB038374	C 894	14	26.9	2318	10	RAT1HCG01	M63918 Rattus norv
C 822	14	26.9	1311	5	HYH224325	C 895	14	26.9	2332	8	SCGYP7	X79552 S.cerevisia
C 823	14	26.9	1323	8	AB021166	C 896	14	26.9	2364	3	TBNNKAI	L03778 Trypanosoma
C 824	14	26.9	1323	8	AY179882	C 897	14	26.9	2478	5	BC045314	BC045314 Danilo rer
C 825	14	26.9	1347	3	AB058667	C 898	14	26.9	2483	8	AOP250266	AJ250266 Amphidini
C 826	14	26.9	1347	3	AB058668	C 899	14	26.9	2496	1	AB084154	AB084154 Paenibacil
C 827	14	26.9	1347	3	AB058669	C 900	14	26.9	2509	8	AF085063	AF085063 Pichia an
C 828	14	26.9	1347	3	AB058670	C 901	14	26.9	2553	8	AK091508	AK091508 Homo sapi
C 829	14	26.9	1347	3	AB058671	C 902	14	26.9	2634	6	AK072900	AK072900 Sequence
C 830	14	26.9	1360	5	AB059657	C 903	14	26.9	2634	6	AX469425	AX469425 Sequence
C 831	14	26.9	1375	8	AB086375	C 904	14	26.9	2634	6	BD080289	BD080289 Isolacion
C 832	14	26.9	1410	6	AR052124	C 905	14	26.9	2638	9	AK026796	AK026796 Homo sapi
C 833	14	26.9	1410	6	AR104310	C 906	14	26.9	2664	5	AF026039	AF026039 Xenopus 1
C 834	14	26.9	1410	6	AR269200	C 907	14	26.9	2742	6	AX653038	AX653038 Sequence
C 835	14	26.9	1410	6	AX079089	C 908	14	26.9	2769	6	AX151623	AX151623 Sequence
C 836	14	26.9	1414	5	BC043981	C 909	14	26.9	2799	6	AX356915	AX356915 Sequence
C 837	14	26.9	1439	5	BC008745	C 910	14	26.9	2799	6	AX356916	AX356916 Sequence
C 838	14	26.9	1459	8	ZMU66404	C 911	14	26.9	2799	6	BD107372	BD107372 Gene HVD1
C 839	14	26.9	1496	6	AX405054	C 912	14	26.9	2814	6	AX714057	AX714057 Sequence
C 840	14	26.9	1510	8	YSCCTH1	C 913	14	26.9	2814	9	AK000685	AK000685 Homo sapi
C 841	14	26.9	1528	3	AK113531	C 914	14	26.9	2814	9	AK056120	AK056120 Homo sapi
C 842	14	26.9	1546	3	AB021864	C 915	14	26.9	2816	8	AY179883	AY179883 Zea mays
C 843	14	26.9	1576	1	RHMA11NTRA	C 916	14	26.9	2826	5	BC045321	BC045321 Danilo rer
C 844	14	26.9	1608	6	AX154567	C 917	14	26.9	2857	5	BC043853	BC043853 Xenopus 1
C 845	14	26.9	1617	6	AX122506	C 918	14	26.9	2861	6	AR072708	AR072708 Sequence
C 846	14	26.9	1617	6	BD164623	C 919	14	26.9	2861	6	E15701	E15701 Human ORP15
C 847	14	26.9	1651	9	AF034838	C 920	14	26.9	2879	8	SCYDL234C	Z74282 S.cerevisia
C 848	14	26.9	1693	6	AX686748	C 921	14	26.9	2902	10	RATLHCG	M26199 Rat luteotopi
C 849	14	26.9	1695	5	AF164114	C 922	14	26.9	2921	8	AY179881	AY179881 Zea mays
C 850	14	26.9	1698	6	AX507943	C 923	14	26.9	2924	3	DMY14171	Y14171 Drosophila
C 851	14	26.9	1715	3	AY051709	C 924	14	26.9	2943	3	AF251294	AF251294 Homo sapi
C 852	14	26.9	1716	1	AB009289	C 925	14	26.9	2960	1	REU279073	AJ7579073 Ralstonia
C 853	14	26.9	1725	6	BC002421	C 926	14	26.9	3078	6	BD012338	BD012338 Genes rei
C 854	14	26.9	1725	6	AX575578	C 927	14	26.9	3078	6	BD158487	BD158487 Primer fo
C 855	14	26.9	1740	6	AX066613	C 928	14	26.9	3123	8	AK027770	AK027770 Homo sapi
C 856	14	26.9	1746	8	AX506758	C 929	14	26.9	3150	6	AX082312	Y07748 O. sativa MR
C 857	14	26.9	1746	8	AY142022	C 930	14	26.9	3163	6	AX405058	AX405058 Sequence
C 858	14	26.9	1764	3	AY069488	C 931	14	26.9	3163	6	AX405058	AX405058 Sequence
C 859	14	26.9	1767	14	HS41ETGA	C 932	14	26.9	3163	10	AY188951	AY188951 Rattus no
C 860	14	26.9	1810	9	BC0097650	C 933	14	26.9	3203	3	AX405050	AX405050 Sequence
C 861	14	26.9	1810	9	BC0098877	C 934	14	26.9	3205	6	AX400031	AX400031 Sequence
C 862	14	26.9	1857	6	AK264389	C 935	14	26.9	3231	1	PSEACCBBC	L14612 Pseudomonas
C 863	14	26.9	1903	9	AK092990	C 936	14	26.9	3250	9	AK095391	AK095391 Homo sapi
C 864	14	26.9	1913	9	BC028158	C 937	14	26.9	3257	6	AX405056	AX405056 Sequence
C 865	14	26.9	1922	9	AK091772	C 938	14	26.9	3277	9	AY188947	AY188947 Homo sapi
C 866	14	26.9	1959	3	LDO430863	C 939	14	26.9	3300	1	AF406639	AF406639 Rhodopseu
C 867	14	26.9	1961	6	BD156517	C 940	14	26.9	3376	8	AB086374	AB086374 Sebandia
C 868	14	26.9	1961	9	AK001634	C 941	14	26.9	3388	1	CBBL5Q	L82000 Rhodobacter

C 942	14	26.9	3452	9	AY188949	AY188949 Homo sapi
C 943	14	26.9	3468	5	AF176086	AF176086 Gallus ga
C 944	14	26.9	3500	6	AR275002	AR275002 Sequence
C 945	14	26.9	3500	6	AX654236	AX654236 Sequence
C 946	14	26.9	3559	1	AB072376	AB072376 SpHingomo
C 947	14	26.9	3573	2	AC019447	AC019447 Drosophila
C 948	14	26.9	3645	5	AF514281	AF514281 Oreochohm
C 949	14	26.9	3689	6	AR206497	AR206497 Sequence
C 950	14	26.9	3689	6	AX147254	AX147254 Sequence
C 951	14	26.9	3741	1	AB091437	AB091437 Birkhoide
C 952	14	26.9	3792	6	AX195972	AX195972 Sequence
C 953	14	26.9	3801	6	AX574106	AX574106 Sequence
C 954	14	26.9	3871	6	BD140451	BD140451 Secretd
C 955	14	26.9	3959	9	AY152396	AY152396 Homo sapi
C 956	14	26.9	3960	8	AF064563	AF064563 Hordomo v
C 957	14	26.9	3962	3	AF145671	AF145671 Drosophill
C 958	14	26.9	3979	6	AR308838	AR308838 Sequence
C 959	14	26.9	4012	8	AY037893	AY037893 Pythophth
C 960	14	26.9	4107	3	AY121618	AY121618 Drosophill
C 961	14	26.9	4180	8	AY0510210	AY0510210 Arabidops
C 962	14	26.9	4283	3	AT0501813	AT0501813 Drosophill
C 963	14	26.9	4404	9	DB7440	DB7440 Human mRNA
C 964	14	26.9	4412	9	BC015052	BC015052 Homo sapi
C 965	14	26.9	4415	6	AX399923	AX399923 Sequence
C 966	14	26.9	4452	9	AK074112	AK074112 Homo sapi
C 967	14	26.9	4462	1	AF126720	AF126720 Mycobacte
C 968	14	26.9	4503	6	AR072702	AR072702 Sequence
C 969	14	26.9	4503	6	E15695	E15695 Human mRNA
C 970	14	26.9	4503	9	HSU65785	HSU65785 Human 150 k
C 971	14	26.9	4620	9	BC033494	BC033494 Homo sapi
C 972	14	26.9	4775	5	GCLR11	Y08109 G.gallus mR
C 973	14	26.9	4810	6	AR206876	AR206876 Sequence
C 974	14	26.9	4810	6	AR308837	AR308837 Sequence
C 975	14	26.9	4810	6	AX367076	AX367076 Sequence
C 976	14	26.9	5163	6	AX308841	AX308841 Sequence
C 977	14	26.9	5228	8	AF072683	AF072683 Emricell1
C 978	14	26.9	5242	7	HSM805174	ALB31879 Homo sapi
C 979	14	26.9	5400	1	AY1HRPZ	L25315 Azotobacter
C 980	14	26.9	5400	14	HSUIJ2	M19120 Herpes simp
C 981	14	26.9	5410	1	AE005715	AE005715 Caulobact
C 982	14	26.9	5451	6	AX078430	AX078430 Sequence
C 983	14	26.9	5471	6	AX305572	AX305572 Sequence
C 984	14	26.9	5471	10	MMNFNG	X025440 Mouse NF-M
C 985	14	26.9	5509	1	BBJ224435	BBJ224435 BiliGobac
C 986	14	26.9	5509	6	AR308836	AR308836 Sequence
C 987	14	26.9	5509	6	AX319625	AX319625 Sequence
C 988	14	26.9	5612	6	AX406977	AX406977 Sequence
C 989	14	26.9	5966	1	SAMY18862	Y18862 Streptomyce
C 990	14	26.9	6099	14	PMO271595	PMO271595 Poinsettia
C 991	14	26.9	6208	6	AX345097	AX345097 Sequence
C 992	14	26.9	6236	6	BD084814	BD084814 Diagnosis
C 993	14	26.9	6236	9	HSNFM	Y0067 Human gene
C 994	14	26.9	6502	6	AX096213	AX096213 Sequence
C 995	14	26.9	6505	9	AB002346	AB002346 Homo sapi
C 996	14	26.9	6505	9	AF039945	AF039945 Homo sapi
C 997	14	26.9	6642	6	I64799	I64799 Sequence 5
C 998	14	26.9	6715	1	B8GALE	X99319 B.subtilis
C 999	14	26.9	6746	9	AF318616	AF318616 Homo sapi
C 1000	14	26.9	6799	9	HSM802436	AL157424 Homo sapi

RESULT 1	CIU39835	1201 bp	mRNA	linear	PLN 18-JAN-1996
LOCUS	CIU39835				
DEFINITION	Coccidioides immitis immunoreactive cell wall protein mRNA,				
ACCESSION	U39835				
VERSION	U39835.1	GI:1161373			
KEYWORDS	Coccidioides posadasii				
SOURCE					

```

ORGANISM      Coccidioides posadasii
REFERENCE     Eukaryote; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Orygenales; mitosporic Orygenales; Coccidioideae.
AUTHORS       1 (bases 1 to 1201)
              Dugger,K.O., Villarreal,K.M., Nguyen,A., Zimmermann,C.R., Law,J.H. and Galgiani,J.N.
TITLE         Cloning and sequence analysis of the cDNA for a protein from Coccidioides immitis with immunogenic potential
JOURNAL       Biochem. Biophys. Res. Commun. 218 (2), 485-489 (1996)
MEDLINE       96144702
PUBMED        8561782
REFERENCE     2 (bases 1 to 1201)
              Dugger,K.O. and Galgiani,J.N.
TITLE         Direct Submission
SUBMITTED     Submitted (01-NOV-1995) Krie O. Dugger, Infectious Diseases, University of Arizona, 1501 N. Campbell, Tucson, AZ 85724, USA
FEATURES      Location/Qualifiers
               source          1..1201
                        /organism="Coccidioides posadasii"
                        /mol_type="mRNA"
                        /strain="Silveira"
                        /db_xref="taxon:199306"
                        152..736
                        /note="33 kDa apoglycoprotein"
                        /codon_start=1
                        /product="immunoreactive spherule cell wall protein"
                        /protein_id="AAC9140.1"
                        /db_xref="GI:1161374"
                        /translation="MOPSHALIVAAAGLASAQLDIPPCALCNFVAIANGDCSTRILTDFCKSKPELPQGITPPCVSEACPLVARISNVIYVDGKACGPIDPVDTTAPAEPSSTIAETPAEPTTEPTAEPTPAETPAETPAETPAETPAETPGTGAVPTGGFTVGPRPTASTPREFPFAGSNVRASVSGIAALAALLGLAYL"

BASE COUNT    249 a           369 c           275 g           308 t

ORIGIN
Query Match   100.0%; Score 52; DB #: Length 1201;
Best Local Similarity 100.0%; Pred. No. 2..7e-17;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy             1 ATGCAGTTCTCTCACGCCCTTCATCGCTGTCTGTCGCTCCGGCTTCGCCACTG 52
                |||||
Db            152 ATGCAGTTCTCTCACGCCCTTCATCGCTGTCTGTCGCTCCGGCTTCGCCACTG 203
                |||||

RESULT 2
LOCUS        CIU32518                   1234 bp mRNA linear PLN 04-NOV-1996
DEFINITION   Coccidioides immitis antigen 2 (Ag2) mRNA, complete cds.
ACCESSION    U32518
VERSION      U32518.1 GI:1200179
KEYWORDS
SOURCE       .
ORGANISM     Coccidioides posadasii
COMMENT      Coccidioides posadasi
AUTHORS      Eukaryote; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Orygenales; mitosporic Orygenales; Coccidioideae.
TITLE        1 (bases 1 to 1234)
              Zhu,Y., Yang,C., Magee,D.M. and Cox,R.A.
SUBJECT      Molecular cloning and characterization of Coccidioides immitis antigen 2 cDNA
JOURNAL
MEDLINE
PUBMED
REFERENCE     2 (bases 1 to 1234)
              Cox,R.A.
TITLE         Direct Submission
SUBMITTED     Submitted (27-JUL-1995) Texas Center for Infectious Disease, Research Immunology, 2303 S.E. Military Drive, San Antonio, TX 78223, USA
On Feb 24, 1996 this sequence version replaced gi:1184058.
location/Qualifiers
               1..1234
                        /organism="Coccidioides posadasii"
                        /mol_type="mRNA"
```


VERSION	AE004517.1	GI:99346710
KEYWORDS	Pseudomonas aeruginosa PAO1	
SOURCE	Pseudomonas aeruginosa PAO1	
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.	
REFERENCE	1 (bases 1 to 10822)	
AUTHORS	Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Mizoguchi, S.D., Warrenner, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltz, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Jim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K.-S., Wu, Z. and Paulsen, I.T.	
TITLE	Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen	
JOURNAL	Nature 406 (6/79), 959-964 (2000)	
MEDLINE	20437337	
PUBMED	10984043	
REFERENCE	2 (bases 1 to 10822)	
AUTHORS	Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Mizoguchi, S.D., Warrenner, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltz, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Jim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K.-S., Wu, Z., Paulsen, I.T., Reizer, J., Sater, M.H., Hancock, R.K.E., Lory, S. and Olson, M.V.	
TITLE	Direct Submission	
JOURNAL	Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA	
REFERENCE	3 (bases 1 to 10822)	
AUTHORS	Pseudomonas aeruginosa Community Annotation Project (PseudocAP) Direct Submission	
CONSRM	Submitted (04-FEB-2003) Department of Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Dr., Burnaby, British Columbia V5A 1S6, Canada	
JOURNAL	This represents the February 3, 2003 version of the continually updated, reviewed, Pseudomonas aeruginosa PAO1 genome annotation, from PseudocAP (see http://www.pseudomonas.com for latest updates and links to alternate annotations). PseudocAP is coordinated by Fiona S.L. Brinkman (Simon Fraser University, Canada) and Robert E.W. Hancock (University of British Columbia, Canada). We welcome submissions through www.pseudomonas.com of any proposed changes.	
COMMENT	'Protein name confidence' is used to rate our confidence of the accuracy of the protein name. Class 1: Function experimentally demonstrated in P. aeruginosa. Class 2: Function of highly similar gene experimentally demonstrated in another organism (and gene context consistent in terms of pathways its involved in, if known). Class 3: Function proposed based on presence of conserved amino acid motif, structural feature or limited sequence similarity to an experimentally studied gene. Class 4: Homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences.	
FEATURES	----- Location/Qualifiers 1. .10822 /organism="Pseudomonas aeruginosa PAO1" /mol_type="genomic DNA" /strain="PAO1" /db_xref="taxon:208964" 81. .506 /locus_tag="PA0817" 81. .506 /locus_tag="PA0817" /note="Protein name confidence: Class 3 (function proposed based on presence of conserved amino acid motif, structural feature or limited sequence similarity to an experimentally studied gene)" /codon_start=1	

gene
CDS

```
/transl_table=1
/product="probable ring-cleaving dioxygenase"
/protein_id="AAG04206.1"
/db_xref="GI:9946711"
/tranlation="MSLMPFHLLAPVHDLPAARFRFYGEFGLSEGRSAEHMVDFFDGG
HOLVHORPOTSDRSHAGSNVDDGHVDVDPHFVGVLAMDWMHALAERLQGRFRVIE
PIVRKGVGQOATPLFLFDPCGNALFEKSPFRMDQLPAK"
complement(525..797)
locus_tag="PA0818"
complement(525..797)
locus_tag="PA0818"
/note="Protein name confidence: Class 4 (homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences)"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAG04207.1"
/db_xref="GI:9946712"
/tranlation="MPYRVILLILALTPALASSATLDERIGDLEHAFFPKPDLOGTSAK
PASPPERFKLAPLGISVYRKKECATSSLPDEOSSCALPECTCAQ"
953..1246
locus_tag="PA0819"
953..1246
locus_tag="PA0819"
/note="Protein name confidence: Class 4 (homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences)"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAG04208.1"
/db_xref="GI:9946713"
/tranlation="MAPRCQVTYAQSPTFIPLRTSCONHPGCATLYLDSSASE
DLFPDRRLTESALGITLSVDLDELTEEPGSTCRITGISLSTPEAYARRRR"
1545..2357
locus_tag="PA0820"
1545..2357
locus_tag="PA0820"
/note="Protein name confidence: Class 4 (homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences)"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAG04209.1"
/db_xref="GI:9946714"
/tranlation="MPAVIAENVDISTWDETGAIVHYPRRYSAILIRGVQVYYKKKL
TKSAPPDRSLASOPHIFGARIGIKVTVIPDPSKDDPALIESFTPEQVALAKGNSY
LELIPSVASNYSWSGVREIGQVTFDSILARAELKPVQAESWGDRALGFESRTGS
LSTYGVNGVEDPTLRLOAIAIHGVCACAGCFPGKAGYGAFTHVHTTQISEFE
APKANPETDPLFCANCAHVHRRDRTLSVDCLKLVGRVAACAADVIV"
complement(2464..3552)
locus_tag="PA0821"
complement(2464..3552)
locus_tag="PA0821"
/note="Protein name confidence: Class 4 (homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences)"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAG04210.1"
/db_xref="GI:9946715"
/tranlation="MEOPFDLAELAAPHILIEIAGNLIMKSGPEDYIAVILCTARGTL
YFAKAHPILVESICCCPFDEFERLABPHLTLMREEBAQCPLTAAYRDQPREMMGA
MEDBDHSIFCTSGSKSRDAGAMLFVDYIGKRSNOAGHDLSLERSVLPLVOEORPL
DFLOLIDFAPLARLEPEGVAGHAVYNISPTSNDKDESPAAMARKGDLVGTACLIAN
TPREKPRIRKIVXSKULTLNNEBRLAAGGDALAQAQPPSSHFAFIRYIGDEVYIOAGATP
TIADDAEBSRAPPYLNLHNAHLKGRIETIOSLHGSGSHDGELRLVGAADQMPLRLDVE
DSEIAPRCDCSLNLEPILDATNTLPERL"
complement(3569..4015)
locus_tag="PA0822"
```

gene
CDS

LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
AE012448/c	Xanthomonas campestris pv. campestris str. ATCC 33913, section 356 of 460 of the complete genome.	AE012448	AE012448.1 GI:21114551	Xanthomonas campestris pv. campestris str. ATCC 33913 Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas.		
LOCUS	131170 bp DNA linear BCT 23-MAY-2002					
ACCESSION	AE012448					
VERSION	AE012448					
KEYWORDS	AE012448.1 GI:21114551					
SOURCE	Xanthomonas campestris pv. campestris str. ATCC 33913					
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;					
REFERENCE	Xanthomonadaceae; Xanthomonas.					
	1 (bases 1 to 131170)					

AUTHORS	TITLE	JOURNAL	PUBLISHED	REFERENCE	AUTHORS
da Silva, A.C.R., Ferro, J.A., Reinach, F.C., Farah, C.S., Furlan, L.R., Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida Jr., N.F., Alves, L.M.C., do Amaral, A.B., Bertolini, M.C., Camargo, L.E.A., Camarotte, G., Canavann, F., Cardoso, J., Chambergo, F., Ciapina, L.P., Ciccarelli, R.M.B., Coutinho, L.L., Cursino-Santos, J.R., El-Dorry, H., Faria, J.B., Ferreira, A.J.S., Ferreira, R.C.C., Ferro, M.I.T., Fornighieri, E.F., Franco, M.C., Greggio, C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite Jr., R.P., Lemos, E.G.M., Lemos, M.V.F., Locali, E.C., Machado, M.A., Madeira, A.M.B.N., Martinez-Bossi, N.M., Martins, E.C., Meidanis, J., Menck, C.F.M., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T.M., Okura, V.K., Oliveira, M.C., Oliveira, V.R., Pereira Jr., H.A., Rossi, A., Sena, J.A.D., Silva, C., de Souza, R.F., Spinola, L.A.F., Takita, M.A., Tamura, R.E., Teixeira, E.C., Tezara, R.I.D., Trindade dos Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C. and Kitajima, J.P.	Comparison of the genomes of two <i>Xanthomonas</i> pathogens with differing host specificities	Nature 417 (6887), 459-463 (2002)	2 (bases 1 to 13170)	da Silva, A.C.R., Ferro, J.A., Reinach, F.C., Farah, C.S., Furlan, L.R., Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida Jr., N.F., Alves, L.M.C., do Amaral, A.B., Bertolini, M.C., Camargo, L.E.A., Camarotte, G., Canavann, F., Cardoso, J., Chambergo, F., Ciapina, L.P., Ciccarelli, R.M.B., Coutinho, L.L., Cursino-Santos, J.R., El-Dorry, H., Faria, J.B., Ferreira, A.J.S., Ferreira, R.C.C., Ferro, M.I.T., Fornighieri, E.F., Franco, M.C., Greggio, C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite Jr., R.P., Lemos, E.G.M., Lemos, M.V.F., Locali, E.C., Machado, M.A., Madeira, A.M.B.N., Martinez-Bossi, N.M., Martins, E.C., Meidanis, J., Menck, C.F.M., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T.M., Okura, V.K., Oliveira, M.C., Oliveira, V.R., Pereira Jr., H.A., Rossi, A., Sena, J.A.D., Silva, C., de Souza, R.F., Spinola, L.A.F., Takita, M.A., Tamura, R.E., Teixeira, E.C., Tezara, R.I.D., Trindade dos Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C. and Kitajima, J.P.	Submitted (28-NOV-2001) Departamento de Bioquímica, Universidade de São Paulo, Av. Prof. Lineu Prestes 748, São Paulo, SP 05508-900, Brazil
FEATURES	TITLE	JOURNAL	PUBLISHED	REFERENCE	AUTHORS
source	Direct Submission				
Location/Qualifiers					
1. .13170					
/organism="Xanthomonas campestris pv. campestris str. ATCC 33913"					
/mol_type="genomic DNA"					
/strain="ATCC 33913"					
/db_xref="ATCC:33913"					
/db_xref="taxon:190485"					
/note="pathovar: campestris"					
118. .1494					
/gene="XCC3308"					
118. .1494					
/gene="XCC3308"					
/note="putative; ORF located using Glimmer/Genemark"					
/codon_start=1					
/cramel_table=11					
/product="hypothetical protein"					
/protein_id="AA042578.1"					
/db_xref="GI:21114552"					
/cramel_table="MSLTSQOQVAAALAKDAVDEPPFGASSRPVDIGITVYKRLVSDS BSGQGIIVRIIDNBIIVAHGCTETRLKQDGYTGGCGAAAHNRQAAELITK HALVYAKLGRGETDVAVTHTSLSGNLAOVTAHFSLDKGTFAVYAGVNSLDRRIPE GGDIVHNWAGDAVASAKIHGVAKLVAQOELATLEKAGYDNRSLVDVNNP KLGDSIRHNNFLPEDANGKPRSVEYDPAQOLAKFAPMIDKRYDDEVLGSGVTLG VGVNGVADGIDALRGTLAPAGSAEMWAAQWQVORIQNDHAPATVVAPEKMPI TSGSEVAPMAPRGPGRGPASVSNPDIYAHISKILPGTSPAPAMVATYKAPI LRADLOSVTSHQDNVWVIVGQIPGRVRYKVLSDSVPLQESIRSQALDTRQSEPT QOSPARVW"					
1499. .1858					
/gene="XCC3309"					
1499. .1858					

```

/gene="XCC3309"
/note="putative; ORF located using Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA042579.1"
/db_xref="GI:21114553"
/translation="MTERVAPPTAPVTVTAARCPPRATSRLOPLADANLGDVFPEKL
VNDYFNREIQDFGMLMERMRWSTIALTTVTLLWMIGYRIATGQSRRESAMATMIK
AGKVAIIISISCEGCRS"
/complement(1861..2646)
/gene="XCC3310"
/complement(1861..2646)
/gene="XCC3310"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA042580.1"
/db_xref="GI:21114554"
/translation="MHGANVAGRLPAPCHAAFSGLKVVFNANGAGAFLEKPMKYLRLH
SPGLAVFALMHPVBARAASWVDAGVTPPGOCOLEAMWRATPGASATVAPCNL
WTEFALGMDTVRHGSAALDIGAKMLRDEQHANGALVSLORAMODPOORA5YIA
VPLTSLGTAATRIHVNRCALVTNAGPDTTACGLSEHVLTPANTVLAASGDDRDG
HCAQJGLRLVACASSIDLTVGRDHATAPRGMTVGINIAMP"
/complement(2963..3469)
/gene="XCC3311"
/complement(2963..3469)
/gene="XCC3311"
/note="putative; ORF located using Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA042581.1"
/db_xref="GI:21114555"
/translation="MLKSLMTMTGILLASPVFAAPAPMORAPGTAA0IGAHTLRO
IPEACALEGVFTGNAOPIRFSYTRSSCTCPRARFVDTYAAASVAGMIFNDYIR
VPSACPSQAVAVRVRKFAAKPOLDGQGSRTYDEAKQ0AAAGKLPQVPMPTAQ0
TMEGRACQ"
3719..5746
/gene="XCC3312"
3719..5746
/gene="XCC3312"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="H+ translocating pyrophosphate synthase"
/protein_id="AA042582.1"
/db_xref="GI:21114556"
/translation="MDEHTGLMLALGCAVLAIVGMLSRWVVAQPSGNARNOEIAAA
IOEGARAYLNRQYLTISAGVLFVLVGFSLWYTAIGFAIGAVLSGLAGYIGNVSV
RANVRTAQARHGISKAMDVARFGAITMLVAGLGLGAGYFALLOMGLPLEONL
HALVGLAFSSLSIFARLGGIFTRGADVADLVGKAGIPEDDPRNPRAVIANDVG
DNVGCAGMAADLFETVATVATLMTLSGLTAAEGPHAVLPLVLCGVSIASTIGA
AFVIRPGSGIMGALYKGVISGVLAATLDELITQMLRENAHGAMSLYACLIGLVL
TGLIVITERTIGTQIPVQVHAASITGHGNTIAGLSISKSTALPVIATVCAAIWG
AFHFGLVGTIAAATAMLSMAGMIVALDVGRIITNAGGIAEMAELEPVRVNTDPLD
AVGNTTKAVTKGAYLGSAAIALVLFAVDTNMLQANPEVAFPLSDHTVIGLIG
GLIPYLFGAMAMEAVGAAAGVAVEERGFRIPIGMTSTAPROYDRAVMDLTSRAIR
EMIVPSLIPVVVAVVIGLLGPRALGGLIGLITVGLFAISMTTGGAAMDAAKRYIE
DGHFGKSEAKKAITDPTVGDPPYKDTAGRAINPLIKINIVALLVPLL"
complement(6144..6626)
/gene="XCC3313"
complement(6144..6626)
/gene="XCC3313"
/note="putative; ORF located using Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA042583.1"
/db_xref="GI:21114557"

```

```

/translation="MSRYLRYIVLIMSSITLAPALAADPPPAFVDAVDMFANGECHEA
FVDLEQRDLDRDFDNI CGDTFCGGEFSFQPLRFCSVNRVSGVNSCMTFCASEVS
DPRSGHLRSDSRWRCRTAPLKAGTRLDEMRYRLAVTNPLFEBLPGCAPPIYNGLICGL
"
6826..7362
/gene="ppa"
/note="synonym: XCC3314"
6826..7362
/gene="ppa"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="inorganic pyrophosphatase"
/protein_id="AA042584.1"
/db_xref="GI:21114558"
/translation="MGLLVYSGKRLPEINIVVIEIPKDEPVEKVEVDKASGAFVDR
ILSTPMRYPCHGVYPTNLGCGDPAVDLVVLPPLVRSVYRCRPVGLRNSDEGS
DEKILAVIERKIFSGYAHIEDIDVSSHMERITGHFPHYKOLEKGMVKLDMGWSGA
EAKRILVESVERVNSDAP"
7509..8621
/gene="XCC3315"
7509..8621
/gene="XCC3315"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="response regulator"
/protein_id="AA042585.1"
/db_xref="GI:21114559"
/translation="WRILFVGGDEASLPDLVDYVIAELDNMQAOQVPDGN5AIEAAL
SPFDVAVIAPVLPDLTAATLGLQIRTLRPDTRIALDAODQRTPPAIIQVAFRL
PMLPAVEVLVPLVLEELRDLISNPRRAIGRIEXKPSPHLYLSLMALEEDDGA
DAADIAKLIAGPATIAAVLQICNSAFSGSGSITDLTAVRLGATVRLDVLASEV
FSVOTLSRAEARMORALLSSRLAAKULPRTSALGSAALAIADIGLLPVRBRE
PVEDEDLGHTGACAYLGLGMLPPIIEAVAFRRHFORSLRSFWTGAHVATAL
ASGETVDEEYLTQGVINRLPWRQADLTMLGLAA"
complement(9251..12142)
/gene="btub"
/note="synonym: XCC3316"
complement(9251..12142)
/gene="btub"

```

```

Query Match      36.5%; Score 19; DB 1; Length 13170;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 27 TTTCTGCTGCTGCCGCCCTC 45
DB 7138 TCTCGTCCGTCGCCGCCCTC 7120

```

```

RESULT 7
AF416330/c
LOCUS AF416330 76093 bp DNA circular BCT 15-MAY-2003
DEFINITION Ruegeria sp. PR1b plasmid pSD20, complete sequence.
ACCESSION AF416330
VERSION AF416330.1 GI:22726324
KEYWORDS
SOURCE
ORGANISM Ruegeria sp. PR1b
Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacteriales;
Rhodobacteraceae; Ruegeria.
1 (bases 1 to 76093)
Zhong Z., Caspi R., Helinski D., Knauf V., Sykes S., O'Byrne, C.,
Shea T.P., Wilkinson, J.E., Deloughery, C. and Toukdarian, A.,
Nucleotide sequence based characterizations of two cryptic plasmids
from the marine bacterium Ruegeria isolate PR1b
Plasmid 49 (3), 233-252 (2003)
2 (bases 1 to 76093)
Zhong Z., Toukdarian, A. and Helinski, D.R.
Direct Submission

```

JOURNAL Submitted (31-AUG-2001) Department of Biology, University of
California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0322,
USA

FEATURES
source Location/Qualifiers

1..76093
/organism="Ruegeria sp. Prib"
/mol_type="genomic DNA"
/strain="Prib"
/db_xref="taxon:185588"
/plasmid="pSD20"
complement(75..1079)
/note="replication protein RepA"
/codon_start=1
/transl_table=11
/product="RB100"
/protein_id="AA05121.1"
/db_xref="GI:22726325"
/translation="MTGISTAAAGYPPEOGAYFLCDIFDAMPKNDLASMEHPLSLG
TRPDRILNENCAEITVTPSVKGLATHDKDLIFCVSOLMALNAGROVSRFLHL
TAHDLVACNRETSQDAYRLREAFERLAGRTITNITTGHEVTSFGILIESMIVR
KARGRWISVAUTLSDMLYRAVLSKSVLISDYPFLAKPLERRIYELARKKCGQDS
WQISVETLLKSGSASPRVRKMRDMIOSPLPDYDISELGDILFPAQGRLVES
APPTVSEAREARRLMPGMVDVHALEADWRAVMARTGAPRLKPPDAFLGWVKKRAOD
"

CDS

repeat_unit
1179..1193
/note="iteron1"
repeat_unit
1201..1215
/note="iteron2"
repeat_unit
1223..1237
/note="iteron3"
1546..1555
/note="replication protein RepA"
/codon_start=1
/transl_table=11
/product="RB101"
/protein_id="AA05122.1"
/db_xref="GI:22726326"
/translation="MAKEIHSKAGPALPYFNIDDALAELODPTARFAEIAEGC
ARGADLASRGINEGSKSLRFSWEITRYLIPVAQAFRRVLKANPDLPOGRSETE
GAKRFTIDEVLRALFAFPAQSGSKANVLPYRPELAKVLAVANFGKGVGTSTAAH
LAMSALDGYKVLVIDDSQSWTSIFGKVEDEWQAFPLMARHKGELRLNQRRL
LDGAPQPLDESLTAMEMTADQVDSIHWPNIDLIGAQLMYAEFOIPVWRMARS
WKLMDALTRLEADQVLDQYDVFIDTPPALGYLTINGLSADILLVPGASFLFDS
TGRFEDMHSFTASIEEGENLAALAGREIGEFDMDAVRAVITRRYDSAOQCELAIVQ
AYMGPTLSPHRODYTALIGQAGEOVNGIYEADYRDPNRETTAKRGSETDATYAAKRL
LTGWRREELAQERRAS"

CDS

CDS

3022..4191
/note="replication protein RepB"
/codon_start=1
/transl_table=11
/product="RB102"
/protein_id="AA05123.1"
/db_xref="GI:22726327"
/translation="MAKRRRLTPAQPTLYLSAERAATPSPAKSGSVGAAPIAQ
VASDTAAALAEELSGVENARAGIMITEIPLEAIDETVLRDVONNEEDMLAMA
SIHSGOOTPIEVVALLEQPOGSHGLISGMRLTALRLAETPAVTYKALVTP
ASADAYLSWEEBENIRVNLISHERARIAMRALEGVFPDOKKALIGLVNSPKRS
KISFVTLVSALDEALREPTALPEKLGIGLAHLEKDPSEFHLERLRASDPADAVS
EMTVNGMTARNVNLNRLERLESQKVPVPRVQPAVTPQAAAPSVQMPRSQAA
APLVVPASSSDRVMLNLAIPGLRLSFAEEAGTIELSGEAVTADMTLKLWLSRR"
4513..5487
/note="putative galactosyltransferase"
/codon_start=1
/transl_table=11
/product="RB103"
/protein_id="AA05124.1"
/db_xref="GI:22726328"
/translation="WTSRKLVAVVVTYNRLDKLTKYTLARLSPAEELAQVVVNNAS
TDGTADMETOHDPRLVLTSETNSGAGFAMGMRHAVETCAPMLVVDQDPAP
GALAFAHLPEDKMLIAAAYFSPSGEICEMNRPSPNPFMLSEFRTALGSGRSGFH
LDPAEYOGCSTIDYTSFVGFISRAIARICYPPPSPIYIGDDSIYTLGLTAAGRI
GFERGVREHDLTFSQEGCAVORGRLTLPWKYTYHNLTLIRLAAGLWMPVLLLI
LPLWLKSRHYGRDGRIFRLLCAVNDGLRRRTSVSEFTVAKWSGEG"

CDS

CDS

5532..7814
/note="putative polysaccharide export protein"
/codon_start=1
/transl_table=11
/product="RB104"
/protein_id="AA05125.1"
/db_xref="GI:22726329"
/translation="MIAFHONPOCQDORODDIERFDIGRIGRIDPCTEADMD
OGIYFDVYLRGCEILVVDHPRGCKQPVODAPSGVBERKVDVDTTRPRPAQSTYH
GKPDRLDHGNVNDRMGVFDQONPANNENAGDKNLDTVEDETRQDVRDGDGRVQER
GRERNPGGVGOAGIKREQIACVCEQKTAQORGAHDREPOROKDGGORLODG
PDIFAGPAPSVRHLSDQADRDRTQGSASMSDLGDGRGCGLLRRKRGCGGHEA
SGEYCVVIRRAGDQTLKMGMTKTANRPGQADMDQDQGLSEATSGDTPAOGKAD
AENPGRGCGKGGGLGPIKAOKLREQVRAEAEMTRHLOGLDOPAPVAPATLK
RRHNGILVSPIYULPLVLYVYILMAIEDOYVSTSGFVARSQATASEALGLAQ
LVGDITASDSQVLAFTQSQEIVEVEREVGIRAFSAHPQDMAFALPDPATLDELIT
WQORILIGISYESGSGMEVSAVDPETAQITAI11AASQDRINALNEQARDAMR
YARADLDVAQOLKAAREALTOFRTRTLDPQADIQCRMGHNLQOOLALALVEYD
LLRGSNDRLKTAETKTIQVIRERIDIEROTFASDNETGGLGENTPSLISEVERL
ITDOYAVOYRASLATELVARDEATRSRLATYIKTLPLESSQYPPRPVAGIAGL
LLLWMSIFVLVYISIRDS"
7827..8486
/note="polysaccharide export protein"
/codon_start=1
/transl_table=11
/product="RB105"
/protein_id="AA05126.1"
/db_xref="GI:22726330"
/translation="MIREPLTKYFMVDGORTVIDNLTLPSSGRSLALGRNAGK
STLEIIGIORPDLGRVYDTGMSVPGLSAFHKDLSGVNRFIRIYGVDTDL
VAFQNFSELAGAFAFSPMRTYSGMSRSLTGASGSIKEDTYLDEVTAVDGKSRDK
SRVIFAEKMTSSAIMVHNLNEVRFQFCAGIVLERGATYFDDIDEALHQAAMMO
G"

CDS

CDS

complement(8493..10283)
/note="no homolog identified by sequence similarity;
hypothetical protein"
/codon_start=1
/transl_table=11
/product="RB106"
/protein_id="AA05127.1"
/db_xref="GI:22726331"
/translation="MITLOSFLTTPDEICSEQALYAHKGAAYVSSQSRQSYMERGCE
VRLDTYVNLFSLDITMARGCQDGLCAEIVGGGRGLGVYHMHGEGSTEFVSDVLTLT
PEQPAQDLDLHLAGNSGGLYLRVYMAADGTVEVSGFATRSLEADRLPRLTISIT
TFABETEVRTERLETPLAEBATRDQIRVOQVNDGAGAGCAQOLESKALATVADPN
LGAGGFARGLLEAEAGASHCLFMDDDASFMENIARTYIFLAHADPATVAAQMI
INTHKMRMESGAMFDCGCRPLFNGVDREBEVEIEMLOAPHIRPLTYGCMWFAP
PIKAVTRHPPEFVAGDDINFLANFNIRTLNGVSDODDTEKESQTLVLDLRH
LVQHLTIPELKRSLGATAKVAIYFTLRSLRYHSTAAVOLMAQDWKMGQPFENI
DMSERRATIKEMSADAEWQPIAQNULSHRGCVAIWSHRARHYKGLFTLNGHILLCCG
ASKLVQDINRRALVIRAFGAARVSTIINRDCAIYVARDRRFFTLMMQAKLTWQFR
RDFPSLQNSYAGYESTRTKGFWORILGKA"
complement(11451..12566)
/note="no homolog identified by sequence similarity; ORF"
/codon_start=1
/transl_table=11
/product="RB107"
/protein_id="AA05128.1"
/db_xref="GI:22726332"
/translation="MAQHKTLIFHLGDKTGSTS IOTAFAOERVOLODASYVPADLA
MNAARNSFAHEAKTPGDKAKAEARIAFARLHQADSPDFIVISADEPENVAVSFO
ELYSRYFGIADBELRLISYVPHARLISGYSERKLCVPAAMDMSMETFVERQDRK
LFTFOFRPAARDAGFDAFQURPMI PAOLHGSVVQDVLHAFGLPTKLIGDEANQ
SLCLEDLMLRLKLVQGRHLTVSADVRHTIGMDIMHKAADMPANPATKRLHSLANDIR
ATYLEDARLDKTEFGKAPLMERELKAIKNAADPPQLSEPSDHLAPELSRLRLAK
VINGLQEGVDPALALHDKVEDIKONRALQPSTEGV"
12776..13981
/note="putative UDP-galactopyranose mutase"
/codon_start=1
/transl_table=11
/product="RB108"
/protein_id="AA05129.1"
/db_xref="GI:22726333"

CDS

CDS

complement(11451..12566)
/note="no homolog identified by sequence similarity; ORF"
/codon_start=1
/transl_table=11
/product="RB107"
/protein_id="AA05128.1"
/db_xref="GI:22726332"
/translation="MAQHKTLIFHLGDKTGSTS IOTAFAOERVOLODASYVPADLA
MNAARNSFAHEAKTPGDKAKAEARIAFARLHQADSPDFIVISADEPENVAVSFO
ELYSRYFGIADBELRLISYVPHARLISGYSERKLCVPAAMDMSMETFVERQDRK
LFTFOFRPAARDAGFDAFQURPMI PAOLHGSVVQDVLHAFGLPTKLIGDEANQ
SLCLEDLMLRLKLVQGRHLTVSADVRHTIGMDIMHKAADMPANPATKRLHSLANDIR
ATYLEDARLDKTEFGKAPLMERELKAIKNAADPPQLSEPSDHLAPELSRLRLAK
VINGLQEGVDPALALHDKVEDIKONRALQPSTEGV"
12776..13981
/note="putative UDP-galactopyranose mutase"
/codon_start=1
/transl_table=11
/product="RB108"
/protein_id="AA05129.1"
/db_xref="GI:22726333"

CDS

CDS

/translation="MRHSRIEQAAMAAOKLIMVAGSGAVIGHHLEAGHDITIVD
ARDHIGNCFTSDSDTGVVH1YGP8H1FHTDAEVMVNFEPFLPKYKNTTSL
DNGQGVNFSLEPNLHTTQFEGTILRPEAHDFITKEADSTIDPQFEEQARFV
GNDLAEAFKGTIKOMGVHPSLEILKRLPVRENDYDFPKHFGQMPENGATSM
IGKILDHGITVOLTSTSPREMGEGVHVFAGLADGVYDLPFGYATLDPFETH
OGDYGCAMVNGESEVPRTITRHKHSPMEHSGSVYRFRSLAEEDDIPYPIR
OQERLALADYKLABETSGVTFVGLGTYRKYLDMDVTIREALDTRGFEAKTAAGDP
IPPFEPAPL"
14006..15517
/note="no homolog identified by sequence similarity;
hypothetical protein"

CDS

/protein_id="AA05130.1"
/product="Rb109"
/db_xref="GI:22726334"
/translation="MVPCVENTETRVVMSHSSASTARQAVVVGMRSGTSALAG
VIARLGCALPEQIMPNANDPNPKGFYESLCAVNLNDGLRLHADTWRPPLADNDS
KIGQELDGTMLAOEYGDAPFLVKDRICILLPYWRLEAEIAPFYLHTRSP
VEVARSLEKRGMPVRLGLLWRYLELEAQTGQRAFTSYDOLQDMGALADIQ
TRCGLDLPEVTEPEICEVEVEGFLSADLRHSTALVCDGSLIQWIGCAQVIGRAAE
GDEPGRALDOIRKAFDAMPFLGASAGAEALAAVPAQOGADAEELAREHAAID
KDELLLOKQAEATLOAVMOAKOMEALAEKTRMBAIAEAKTRMEALAEKTR
KQTRRLKMQVEQKAYEDLSTALSGORHAKMGEELRAEMQRAVVARDDAQALV
HDEETERLRRLERQIIDREVLQERQALRQSTWRLLTAPLRMNVNGLGRA"
15514..19611
/note="putative glycosyl transferase"
/codon_start=1
/transl_table=11
/product="Rb110"
/protein_id="AA05131.1"
/db_xref="GI:22726335"
/translation="MISGSGESMTTHADLSLFOSDWYLEQNPVADHFGVSALEH
YHFGMRGERSPHPLFVAMVYLSQSPESCEGLPLTHLFGMKMRSPHVLNPPAA
YLRNDDLOADINPMLYYLQSGMOGEGRLPDRMYLAVYEDVAGIGIPELSYLS
RGMREGRAHPLFASGWAERFQOLGADENPMLHYTTTQWREMRPHRTFDRKWHDS

Query Match 36.5%; Score 19; DB 1; Length 76093;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 GTTCGCGCGCGCTGCCCA 49
Db 49047 GTTCGCGCGCGCTGCCCA 49029

RESULT 8
AP005679/c 134963 bp DNA linear HTG 06-SEP-2002
LOCUS Oryza sativa (japonica cultivar-group) chromosome 9 clone
DEFINITION OJ1210_A07, *** SEQUENCING IN PROGRESS ***
ACCESSION AP005679
VERSION HTG: HTGS PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1
Sasaki, T., Matsumoto, T. and Katayose, Y.,
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, BAC
clone: OJ1210_A07
Published Only in Database (2002)
2 (bases 1 to 134963)
Sasaki, T., Matsumoto, T. and Katayose, Y.,
Direct Submission
Submitted (05-SEP-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasakik@nias.affrc.go.jp, URL: http://rpg.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
The nucleotide sequence of this BAC clone was generated by

FEATURES

source

1..134963

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="9"

/clone="OJ1210_A07"

BASE COUNT 36672 a 30063 c 30068 g 37998 t 362 others

ORIGIN

Query Match 36.5%; Score 19; DB 2; Length 134963;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 CTCATCGCTCTCGTGGCTG 37
Db 76439 CTCATCGCTCTCGTGGCTG 76421

RESULT 9
AP003250/c 146670 bp DNA linear PLN 29-MAR-2003
LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
DEFINITION PAC clone: P0443D08.
ACCESSION AP003250
VERSION AP003250 BA000010
KEYWORDS AP003250.3 GI:15408719
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1
Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T.,
Katayose, Y., Wu, J., Mimura, Y., Cheng, Z., Nagamura, Y.,
Antonio, B.A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K.,
Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K.,
Hamada, M., Harada, C., Hijiishi, S., Honda, M., Ichikawa, Y.,
Idonuma, A., Iijima, M., Ikeda, M., Ikono, M., Itoh, S., Itoh, T.,
Itoh, Y., Itoh, Y., Iwabuchi, A., Kamuya, K., Karasawa, W., Katagiri, S.,
Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Mehara, T.,
Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashina, M.,
Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M.,
Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T.,
Shomura, A., Song, J., Takazaki, Y., Teresawa, K., Tsuji, K., Waki, K.,
Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K.,
Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J.H., Kim, H.I., Eun, M.Y.,
Yano, M., Jiang, J. and Gojobori, T.
The genome sequence and structure of rice chromosome 1
Nature 420 (6913), 312-316 (2002)
22337376
12447438
2 (bases 1 to 146670)
Sasaki, T., Matsumoto, T. and Yamamoto, K.,
Direct Submission
Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasakik@nias.affrc.go.jp, URL: http://rpg.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Aug 31, 2001 this sequence version replaced gi:14522861.

Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI Nonredundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RCP. Protein homologies of the coding regions were searched against NCBI Nonredundant Protein database with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DBJ accession no. and RCP clone ID. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from T7 to SP6 of the PAC clone. This sequence of P0443D08 clone has an overlap with OSJNBa0083M16 (DBJ: AP003214) at the position 1 to 624 and with P0480E02 (DBJ: AP002913) clone at the position 141,467 to 146,670 of 3' end. The sequence of this clone starts at the position 138,088 of OSJNBa0083M16 and ends at the position 5205 of P0480E02. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rhp.dna.affrc.go.jp/Genomeseq.html>.

FEATURES

source

```

1. 146670
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0443D08"
2727..3227
/gene="P0443D08.1"
2727..3227
/note="contains ESTs D48568 (S14863), AU089765 (S14863)
similar to 16.9 kDa heat shock protein
unknown protein"
/codon_start=1
/protein_id="BAB64123.1"
/db_xref="GI:15408720"
/translation="MSLVKLEDTLAFDANPFPS1FGTTVAADAMLASDTSNANTYLE
SRETAAYVFRADLPAGVKEVEVEDEGNLVITGERSVREKORSIHRSICA
TFGRFHLDDAVDLVRASMDGMLTVTVPKVTIDQPAIAAAPPVAVAPAVEAK
AIEASP"
complement(join(4071..4126,4283..4398,5525..5601))
/gene="P0443D08.2"
complement(join(4071..4126,4283..4398,5525..5601))
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64124.1"
/db_xref="GI:15408721"
/translation="MSYAMFLAKINPKFDNAKFFRRVCARPOSTAGKACKXLRHHGEY
AGIQSVNRSPLMLSLFPGCGVGLLSATVRRYLTH"
complement(6051..6500)
/gene="P0443D08.3"
complement(6051..6500)
/note="P0443D08.3"
/note="contains EST AU093701 (C63333)"
/codon_start=1
/product="putative 16.9 kDa heat shock protein"
/protein_id="BAB64125.1"
/db_xref="GI:15408722"
/translation="MSLVRSNVFDPDFADWDPDGVFRSLVPAATSDRTAAAFANARV
DKKETPESHVFKADLPGVKKEVEVEEGNVLVITSGORSKEKEDKDKKHRRVERSSG
QFMRFRRLPENAKVDQVKAEMGVLTVTVPKAEVKKPEVKAIEISC"
complement(8660..10035)
/gene="P0443D08.4"
complement(8660..10035)
/gene="P0443D08.4"

```

gene
CDS

/note="pseudogene
similar to non-LTR retroelement reverse transcriptase"
/pseudo
complement(10583..11035)
/gene="P0443D08.5"
complement(10583..11035)
/gene="P0443D08.5"
/note="contains EST AU092880 (C60687)"
/codon_start=1
/product="putative 16.9 kDa heat shock protein"
/protein_id="BAB64126.1"
/db_xref="GI:15408723"
/translation="MSLVRSNVFDPFSLDMDPDSVFRSVPAATSDNTAAAFANAR
IDKETPESHVFKADLPGVKKEVEVEEGNVLVITSGORSKEKEDKDKKHRRVERSS
QFMRFRRLPENAKVDQVKAEMGVLTVTVPKAEVKKPEVKAIEISC"
13625..14077
/gene="P0443D08.6"
13625..14077
/gene="P0443D08.6"
/note="contains EST AU092880 (C60687)"
/codon_start=1
/product="putative 16.9 kDa heat shock protein"
/protein_id="BAB64127.1"
/db_xref="GI:15408724"
/translation="MSLVRSNVFDPFSLDMDPDSVFRSVPAATSDNTAAAFANAR
IDKETPESHVFKADLPGVKKEVEVEEGNVLVITSGORSKEKEDKDKKHRRVERSS
QFMRFRRLPENAKVDQVKAEMGVLTVTVPKAEVKKPEVKAIEISC"
complement(join(16992..17439,18474..18547))
/gene="P0443D08.7"
complement(join(16992..17439,18474..18547))
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64128.1"
/db_xref="GI:15408725"
/translation="MEGCGQGWTAHFVDEVEREKRPVKFVULVRLCCRRK
SYRDPVAAAGEEDDEKESLLTSSLEELVDDDDCAIDDAVTCNALSIAK
KDGAAPVVLPPGCLHHPVWAPRPHGCVTTSSGGRCOAQVCCRFRPFGLRRRLKM
RRPRFVLVAIPF"
complement(join(19682..20619,20791..21045,21998..22714,
23306..23926,24506..25205))
/gene="P0443D08.8"
complement(join(19682..20619,20791..21045,21998..22714,
23306..23926,24506..25205))
/gene="P0443D08.8"
/note="contains ESTs AU030772 (E60216), AU030771 (E60216)
similar to wall-associated kinase
unknown protein"
/codon_start=1
/protein_id="BAB64129.1"
/db_xref="GI:15408726"
/translation="MSPSFFVVVSAWSLALMLAAARGAEEGGGCLSGQKCDLN
ISSPFWIIQOQADKPCGPDLOYVCCNNSTVATLRSDTSDFDIIINSYGRMLVFD
VHKLARLNNSTGCSI PVFNTFAKLPITFISPHNNLVFVCTAPPAEQOOLGVE
TRCGNNTFARLGRFEGGYDYKYLEGCRNSVFLPVEPPGKMAASRYVVLGCG
GFLITMOLPPVLTSSAANTAADNRPOGCASTYCCGVTI SSPAVVPEQATEKCKH
LGQVIVCHNTPFLYLKPKRYOIIDITFGNNSLAVSDIHKADRTVFSVSEYSC
HVERINTSSKRVGLPFSISTNNLFLYSCKNALVPRGDDLVETRCGNKTFARVGN
YSDSGDPAFYMEGCNATVVPVLGTDARSYEOLIRDFGLTWQETPSASFLLEPAPA
SSSGPCLPTPCCKLTISYPMLEERPSPGSPFOLKCNATGATVLTHTIYEARVY
DIPTGNHTVHVDENLPLATGCPAPRNISDGIWGPVFTVSEANAELRFLSCSLAM
AAAPRPHSLPCDDONSVALVSDHILHEGIGPGCAFVTYVYORNGSMAGTIAAM
RSGGILFEMAVSDDCRKQVSGNCTYSDDLFRACNCPDGMHPDKCFRSESHAG
SSGGLILVSIPTFAMHKRRKROTKDLKOLHSSSSMOSYSKDLGSGSPHIFTYE
LEBATAGFSASRELGGOGFGTVYKGLRDGRVVAVKLYKNYRVRVOPFLNEVDILSR
LHQNVILVIGCTSSRSRDLVLEYEYPNGTVADHLGPPAGEGLTVPYMTJAITET
AEALAYLHAEIIRHDVKTNNILLDNHFVKVADFGSLRPLEVETVSTVPOSTPGY
VDVPEYQCVLTQKSDVYSGVVLIELISSKPAVDMRSRSHDINLEAMALNRONHEV
DOLVDEIYEYETSETKRWVDVLAELAFQCLQMDRESRPITKEVETVLNCKIGCECA
EKKNKASPKEDSHLKDLSQYSPDVIYTHFQSQSTHSAVSSSG-
complement(join(27648..27664,28123..28903))
/gene="P0443D08.9"
complement(join(27648..27664,28123..28903))

gene

CDS

gene

gene

CDS

gene

gene

CDS

gene

gene

CDS

gene

gene

CDS

gene

gene

CDS

gene

gene

CDS

gene

gene

CDS

gene

gene

CDS

gene

gene

CDS

gene

gene

CDS

gene

/gene="P0443D08.9"
 /note="hypothetical protein"
 /codon_start=1
 /protein_id="BAB64130.1"
 /db_xref="gi:15408727"
 /translation="MHPTLICPLASLLILCHRAAECEPATCGLTRYPFMIGCP
 NFNQSQSPSSALASCGHPAEVWNGVASLRGQILVLISVSSFFVAHKRVA
 DGGDVCRITDFN1SSLSLSPFTISSNRAICFLYSNCTEPEIDGLVNATISSCSK
 PIYAVLGIGYDRDNPPAIKAGNCTYSYLVLWPDSPANLTAGTNVSPQFKGFLVEMQ
 KNKFGCDACNGCGCCRYINDSAAAFACISDGLRSTCGORSE"
 complement(join(29901..29938,29998..30738,30790..30937,
 30978..31158,31206..31362,31389..31517,31689..31800,
 31910..32176))
 /gene="P0443D08.10"

Query Match 36.5%; Score 19; DB 8; Length 146670;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 CGCTCTGTCGTCGCGC 42
 |||||
 Db 68656 CGCTCTGTCGTCGCGC 68638

RESULT 10 156485 bp DNA linear HTG 11-JUL-2002
 AP005509 Oryza sativa (japonica cultivar-group) chromosome 8 clone
 LOCUS OJ1003 A09, *** SEQUENCING IN PROGRESS ***.
 ACCESSION AP005509.1 GI:21728111
 VERSION HTG; HTGS_PHASE2.
 KEYWORDS Oryza sativa (japonica cultivar-group)
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE 1
 AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
 clone: OJ1003 A09
 JOURNAL Published Only in Database (2002)
 REFERENCE 2 (bases 1 to 156485)
 AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.
 TITLE Direct Submission
 JOURNAL Submitted (10-JUL-2002) Takuji Sasaki, National Institute of
 Agricultural Sciences, Rice Genome Research Program; Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/,
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 COMMENT The nucleotide sequence of this BAC clone was generated by
 combining Monsanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 Location/Qualifiers
 1..156485
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="8"
 /clone="OJ1003 A09"

BASE COUNT 43887 a 34663 c 35167 g 42451 t 317 others
 ORIGIN

Query Match 36.5%; Score 19; DB 2; Length 156485;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 ACCGCTCATCGCTCTCGT 32
 |||||
 Db 91526 ACCGCTCATCGCTCTCGT 91508

RESULT 11 157945 bp DNA linear HTG 12-SEP-2002
 AP005742 Oryza sativa (japonica cultivar-group) chromosome 9 clone P0489D11,
 LOCUS AP005742.1 GI:22830820
 DEFINITION *** SEQUENCING IN PROGRESS ***.
 ACCESSION AP005742
 VERSION HTG; HTGS_PHASE2.
 KEYWORDS Oryza sativa (japonica cultivar-group)
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE 1
 AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC
 clone: P0489D11
 JOURNAL Published Only in Database (2002)
 REFERENCE 2 (bases 1 to 157945)
 AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.
 TITLE Direct Submission
 JOURNAL Submitted (11-SEP-2002) Takuji Sasaki, National Institute of
 Agricultural Sciences, Rice Genome Research Program; Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/,
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 COMMENT NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 Location/Qualifiers
 1..157945
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="9"
 /clone="P0489D11"

BASE COUNT 45183 a 35304 c 34626 g 42530 t 302 others

Query Match 36.5%; Score 19; DB 2; Length 157945;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 CTCATGCTCTGTCGCTG 37
 |||||
 Db 2999 CTCATGCTCTGTCGCTG 3017

RESULT 12 172081 bp DNA linear HTG 21-FEB-2002
 AP004776 Oryza sativa (japonica cultivar-group) chromosome 2 clone P0452F04,
 LOCUS AP004776.1 GI:18844996
 DEFINITION *** SEQUENCING IN PROGRESS ***.
 ACCESSION AP004776
 VERSION HTG; HTGS_PHASE2.
 KEYWORDS

SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretiaceae; Oryzae; Oryza.

REFERENCE
AUTHORS
1 Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC
clone: P0452F04
Published Only in Database (2002)

JOURNAL
REFERENCE
AUTHORS
TITLE
2 (bases 1 to 172081)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (20-FEB-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: sasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1. 172081
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="P0452F04"
BASE COUNT 49203 a 36993 c 37880 g 47919 t 86 others
ORIGIN

Query Match 36.5%; Score 19; DB 2; Length 172081;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 CGCTCTCGTCGTCGCCGC 42
|||
Db 121561 CGCTCTCGTCGTCGCCGC 121543

RESULT 13
AP005937 298900 bp DNA linear BCT 28-MAR-2003
LOCUS
DEFINITION
Bradyrhizobium japonicum USDA 110 DNA, complete genome, section
3/31.
ACCESSION
AP005937 BA000040
VERSION
AP005937.1 GI:27348808
KEYWORDS
ORGANISM
Bradyrhizobium japonicum USDA 110
Bradyrhizobium japonicum USDA 110
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Bradyrhizobiaceae; Bradyrhizobium.

REFERENCE
AUTHORS
1 Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T.,
Sasamoto, S., Watanabe, A., Ideawa, K., Iriiguchi, M., Kawashima, K.,
Kohara, M., Matsumoto, M., Shimpou, S., Tsuruoka, H., Wada, T., Yamada, M.
and Tabata, S.
Complete genomic sequence of nitrogen-fixing symbiotic bacterium
Bradyrhizobium japonicum USDA110
JOURNAL
MEDLINE
PUBMED
DNA Res. 9 (6), 189-197 (2002)
22484998
12597275

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Caneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T.,
Sasamoto, S., Watanabe, A., Ideawa, K., Iriiguchi, M., Kawashima, K.,
Kohara, M., Matsumoto, M., Shimpou, S., Tsuruoka, H., Wada, T., Yamada, M.

TITLE
Kohara, M., Matsumoto, M., Shimpou, S., Tsuruoka, H., Wada, T., Yamada, M.
and Tabata, S.
Complete genomic sequence of nitrogen-fixing symbiotic bacterium
Bradyrhizobium japonicum USDA110 (supplement)
JOURNAL
MEDLINE
PUBMED
DNA Res. 9 (6), 225-256 (2002)
12597279

REFERENCE
AUTHORS
3 (bases 1 to 298900)
Kaneko, T.
Direct Submission
Submitted (20-NOV-2002) Takakazu Kaneko, Kazusa DNA Research
Institute, The First Laboratory for Plant Gene Research: 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: kaneko@kazusa.or.jp,
URL: http://www.kazusa.or.jp/rhizobae/,
Tel: 81-438-52-3935 (ex. 2338), Fax: 81-438-52-3934)
Location/Qualifiers
1. 298900
/organism="Bradyrhizobium japonicum USDA 110"
/mol_type="genomic DNA"
/strain="USDA110"
/db_xref="taxon:224911"
/complement (114, .329)
/gene="bs10560"
complement (114, .329)
/note="ORF ID: bs10560
hypothetical protein"
/codon_start=1
/protein_id="BAC45826.1"
/db_xref="GI:27348809"
/translation="MTRFLVYISLLLSAAGSAAQOQGHDAQADVTRFCRPVNNNGD
QAVLACQKQNRTRLSKQCDKVLTEHGO"
complement (495, .818)
/gene="b110561"
complement (495, .818)
/note="ORF ID: b110561
unknown protein"
/codon_start=1
/protein_id="BAC45826.1"
/db_xref="GI:27348810"
/translation="MTKSTFRKSLTSSLSKINDTSLALCFPAIVSVITSSAFS
AEAOQOCTGDAFLCSSEIPVPIKRTACMKHRSDLGACRAVMDKLAKGSRKVD
AQQNQ"
1073, .2128
/gene="b110562"
1073, .2128
/note="ORF ID: b110562
putative polysaccharide deacetylase"
/codon_start=1
/protein_id="BAC45827.1"
/db_xref="GI:27348811"
/translation="MKNALGMLASVAVAIVIAAGWFYSSADPGAPKTVAAARAD
PLPAQRTIAKDDVAITPAIAAKPTAAVAQASAPTPAPVAPQQAQACNPAALGAVR
VEIDTTGPGFGEHFKQDFLTDKEVVLFTDDQPMVPNTPAVLAALDECTKGFFS
VGKATYHPRIEILROVLQOHTGVTGHTSHVNLNLSKMTTEOQVKEVEGFAVPAAG
TNAPRFPRLPOLHNPAMVYVFCGRNVAMFSTIDSTDFRFGATPEKIVETVWVRDQ
LGGIITIMHPFOKHTGSAFAMALRLARLAKGIVKVVQAKTTFOTLPETDEALLDLKV
PTSSARPTGSVVQVSO"
complement (2294, .2662)
/gene="b110563"
complement (2294, .2662)
/note="ORF ID: b110563
hypothetical protein"
/codon_start=1
/protein_id="BAC45828.1"
/protein_id="BAC45828.1"

FEATURES
source

gene
CDS


```
/db_xref="GI:27348812"
/translation="MKRIIVAAVVLASASQAAGPRLSLATSEQASAPQATATTIA
VTVVSEAPVTEAPKVVDPAPISAPAPAAVATTSPAAATATKTAKVKKPKK
KRTWTEGRIISELHRHGIYM"
complement(2830..2905)
/nc="crlnt-cgu"
/gene="b110564"
/complement(3005..4747)
/nc="b110564"
/complement(3005..4747)
/nc="ORF_ID:b110564
hypochemical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC45829.1"
/db_xref="GI:27348813"
/translation="MLRIIVLFVLIALAAGAAVADQGDVLITAGGFRASTILPRF
VEGLFAAAVVLVMSLITTIWRTGRLRRSHKRRHARGHAIHGLIAGHGTAL
ARRHAETRRRHAPNDPLALLHAQSAQLEGRNDEMORAFRAMAREDRILLGLGLFI
EAQRADDAVGAAMIABEAIKLSPSSTASHVLGRCARGMSGALAILDSNGLI
DKPAYRRQGVLLTARALELOTMDVARSEVMEAIKLAPTLVPAVLAARESEAHQ
VRRAMKLVAAAMLNPHDLAEAVYHVKLGDSARQRLQVETLAAKTADVAGHIEGO
LALRAAIDASEFSRARVLAIVYNDPTQVALLMAIEIERTHEGDSGARAMTLBAVR
ARHDPATADGVSDRRRVSVPVGTGLDAFQMTQVASLPSGDLISSAPEEMILA
APPKRVTPAPNDTLMERPAPASVPAEDONSPPAKKEPKKEPMTSAEPYTPAPEA
ESSPAATPVFRTRADLCKPAPVPIPAVPIVRAVDPDGIDEGSDEFTQIGTPKA
QAGGRGFWRMGA"
complement(4784..5995)
/gene="b110565"
/complement(4784..5995)
/nc="b110565"
/nc="ORF_ID:b110565
hypochemical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC45830.1"
/db_xref="GI:27348814"
/translation="MADKPPDAGLAPESGRAPRPPTIDLEATDVSTQPOBAGEPE
APVSEHTEPQAKSEQDAEPEREAQAVAPAPASVPSPVITAPISGVAVAVVI
AVGWMGLPAPVQAPPAAPQVTSATVDALGCVAAVEAAGKPADPAPVAVARIDLEKS
VGLRSDIANLAQSDKTASALNDKASAPGALPDALYDRIOLEPAGKTERAELA
QOGEKIDARTMDKPLRHVVAAALLDVAHGDPEYEQLAARSIAKPEMLKLOT
PASSGIPTPVALSRELINIVSKLSPPAOSSGGIVRLRAGASKLVIRERTGVGN
DRGALVAVTAAALRNDVEARELKTLPDADRAPQAWLDKARADALAASKRFAD
DAMADLVKPAQ"
complement(6085..6831)
/gene="b110566"
/complement(6085..6831)
/nc="b110566"
/nc="ORF_ID:b110566
putative utrophinogen III synthase (EC 4.3.1.8)"
/codon_start=1
/transl_table=11
/protein_id="BAC45831.1"
/db_xref="GI:27348815"
/translation="MSLIVTRPHPDNENATDNLRARGHVVLAPVKEEPAFDESE
ANVSALIVTSANAIRAVAPOLQDGLKLPFPAVCEHTATAPARAGFEVITAGDSA
ALRDTVVSARDKVLKKSSTLYLAGABLSRDEGETEGFNVTQTTVMAPVXVL
PRDVEGFAAHGIEAVLHYSRRSAAFLDARDEGVEISALAIPOCCLSFVASVLRD
AGASQVLAAPKPRDENALFTLERALQTRLA"
6983..8056
/gene="gcp"
/nc="gcp"
/nc="ORF_ID:b110567"
/nc="gcp"
/nc="ORF_ID:b110567"
/codon_start=1
/transl_table=11
/product="O-sialoglycoprotein endopeptidase"
/protein_id="BAC45832.1"
/db_xref="GI:27348816"
/translation="MLVIGIETCTDETAANVTERAPDQIGKILSNIVRSQVDEHAPFG
GVPEIARAHAVDLDDGIIDRAKREAGIGFAQLNGVAAAGPGLIGGIYGLITRAKI
ALVHDTPLVAANHLEAHALTRLDGIEFPYCLFLASGHTQIVAVITGVQYVLTGTT
```

```
VDDAIGCFDVKVAMKLGIPYRGQOVERMAAGGAARFAFRPRPHOGRPDANFSLSGLK
TAVNERSRIAEITPDI SDLCASFQAALVETADRLVNGJLRFREOCGAPRALVAG
GVAANOAI RGALHDVARQAKTOLIMPPALCTDNGAMAMAGABRLALGMDTWEAP
RARMULDANATAPAGYKTRAG"
8245..9225
/gene="gpa"
/nc="gpa"
/nc="number=1.1.1.94"
/nc="ORF_ID:b110568"
/codon_start=1
/transl_table=11
/product="glycerol-3-phosphate dehydrogenase [NAD(P)+]"
/protein_id="BAC45833.1"
/db_xref="GI:27348817"
/translation="MTAFHSVAVIGACAMGTALAMVAAARAGSVTLMARNAEHATRIA
STRNPLPGVRLADTIVVSELLARADMLLATPPOHRLGVVNMASHARTPV
VACKGIEHGCHKMTVEVIAEALHAPALILSGSPADVARGLPATATLAKFEELA
SELVALGSPTRFPHYSTDVRGVEIGGAQVLAIAAGIVVGRNIGASALALTTTRGF
SELRILRACGARETILSGSLDLDLSTSAOSRNFPALGIALGRGEAAPAGKLAEG
AFTAPVVELMAARNVMPVSOAVAILDNKLTIDALIEGLTRPFKAEE"
9344..9790
/gene="b110569"
/nc="b110569"
/nc="ORF_ID:b110569
hypochemical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC45834.1"
/db_xref="GI:27348818"

Query Match 36.5%; Score 19; DB 1; Length 298900;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 TCGTCGCTGCCGCCCTCGC 47
Db 273395 TCGTCGCTGCCGCCCTCGC 273413

RESULT 14
AY142898
LOCUS AY142898 1500 bp DNA linear BCT 19-DEC-2002
DEFINITION Helibacillus mobilis Inosine-5'-monophosphate dehydrogenase gene,
partial cds.
ACCESSION AY142898
VERSION AY142898.1 GI:27262441
KEYWORDS Helibacillus mobilis
SOURCE Helibacillus mobilis
ORGANISM Helibacillus mobilis
REFERENCE 1 (bases 1 to 1500)
AUTHORS Raymond,J., Zhaxybayeva,O., Gogarten,J.P., Gerdes,S.Y. and
Blakenship,R.E.
TITLE Whole-genome analysis of photosynthetic prokaryotes
JOURNALS Science 298 (5598), 1616-1620 (2002)
MEDLINE 22337798
PUBMED 12446909
REFERENCE 2 (bases 1 to 1500)
AUTHORS Liolios,K.G., Chu,L., Ostrovskaya,O., Mendybaeva,N.,
Koukharenko,V., Gerdes,S., Kyripides,N. and Overbeek,R.
TITLE Direct Submission
JOURNAL Submitted (19-AUG-2002) Bioinformatics, Integrated Genomics, Inc.,
2201 W. Campbell Park Dr., Chicago, IL 60612, USA
FEATURES
source
1..1500
/organism="Helibacillus mobilis"
/mol_type="genomic DNA"
/db_xref="taxon:28064"
/nc="RHMO03875"
1..>1500
```


* 2376 2842: contig of 467 bp in length
 * 2843 2922: gap of unknown length
 * 2923 3490: contig of 568 bp in length
 * 3491 3570: gap of unknown length
 * 3571 3985: contig of 415 bp in length
 * 3986 4065: gap of unknown length
 * 4066 4452: contig of 387 bp in length
 * 4453 4533: gap of unknown length
 * 4533 5108: contig of 575 bp in length
 * 5108 5188: gap of unknown length
 * 5188 5875: contig of 688 bp in length
 * 5876 5955: gap of unknown length
 * 5956 6557: contig of 602 bp in length
 * 6558 7066: contig of 429 bp in length
 * 7067 7146: gap of unknown length
 * 7147 7793: contig of 647 bp in length
 * 7794 7873: gap of unknown length
 * 7874 8857: contig of 984 bp in length
 * 8858 8937: gap of unknown length
 * 8938 9508: contig of 571 bp in length
 * 9509 9588: gap of unknown length
 * 9589 10479: contig of 891 bp in length
 * 10480 10559: gap of unknown length
 * 10560 12787: contig of 2228 bp in length
 * 12788 12867: gap of unknown length
 * 12868 14642: contig of 1775 bp in length
 * 14643 14722: gap of unknown length
 * 14723 16281: contig of 1559 bp in length
 * 16282 16361: gap of unknown length
 * 16362 17613: contig of 1252 bp in length
 * 17614 17693: gap of unknown length
 * 17694 19013: contig of 1320 bp in length
 * 19014 19093: gap of unknown length
 * 19094 21117: contig of 2024 bp in length
 * 21118 21197: gap of unknown length
 * 21198 22761: contig of 1564 bp in length
 * 22762 22841: gap of unknown length
 * 22842 25899: contig of 3058 bp in length
 * 25900 25979: gap of unknown length
 * 25980 31846: contig of 5867 bp in length
 * 31847 31926: gap of unknown length
 * 31927 37877: contig of 5951 bp in length
 * 37878 37957: gap of unknown length
 * 37958 44429: contig of 6472 bp in length
 * 44430 44509: gap of unknown length
 * 44510 51326: contig of 6817 bp in length
 * 51327 51406: gap of unknown length
 * 51407 57379: contig of 5973 bp in length
 * 57380 57459: gap of unknown length
 * 57460 58102: contig of 643 bp in length
 * 58103 58182: gap of unknown length
 * 58183 58489: contig of 307 bp in length
 * 58490 58569: gap of unknown length
 * 58570 59053: contig of 484 bp in length
 * 59054 59133: gap of unknown length
 * 59134 59822: contig of 689 bp in length
 * 59823 59902: gap of unknown length
 * 59903 60285: contig of 383 bp in length
 * 60286 60365: gap of unknown length
 * 60366 60953: contig of 588 bp in length
 * 60954 61033: gap of unknown length
 * 61034 61401: contig of 368 bp in length
 * 61402 61481: gap of unknown length
 * 61482 62093: contig of 612 bp in length
 * 62094 62173: gap of unknown length
 * 62174 62583: contig of 410 bp in length
 * 62584 62663: gap of unknown length
 * 62664 62966: contig of 303 bp in length
 * 62967 63046: gap of unknown length
 * 63047 63353: contig of 307 bp in length
 * 63354 63433: gap of unknown length
 * 63434 63781: contig of 348 bp in length.

FEATURES
 source
 Location/Qualifiers
 1..63781
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /strain="y, cn bw sp"
 /db_xref="taxon:7227"
 /chromosome="2"
 /map="41C-41C"
 /clone="BACR04F06 (D1057) RPCI-98 04.F.6"
 /clone_1ib="RPCI-98 (Roswell Park Cancer Institute
 Drosophila melanogaster BAC library, partial EcoRI in
 pBAC3.6"
 BASE COUNT 17682 a 13292 c 12914 g 16693 t 3200 others
 ORIGIN
 Query Match 34.6% Score 18; DB 2; Length 63781;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 CATCGCTCTGCTGCTGC 38
 |||||
 Db 34873 CATCGCTCTGCTGCTGC 34890
 |||||
 RESULT 17
 AC025921
 LOCUS Homo sapiens clone RP11-23F24, LOW-PASS SEQUENCE SAMPLING.
 DEFINITION
 AC025921
 AC025921.1 GI:7259725
 VERSION
 KEYWORDS
 HTG; HTGS PHASE0.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE 1 (bases 1 to 64271)
 JOURNAL Homo sapiens, clone RP11-23F24
 REFERENCE
 AUTHORS 2 (bases 1 to 64271)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Baerlein, V., Bede, F.,
 Boguslavskiy, L., Bouckgalter, B., Brown, A., Burkett, G.,
 Campiolo, A., Casale, A., Choepel, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J.S.,
 Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Govette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
 Howland, J.C., Iliev, T., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J.,
 Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPeeters, R.,
 Meldrum, J., Menues, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testafave, S., Theodore, J., Titrill, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viet, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (17-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information

Center project name: L4356
Center Clone name: 23_F_24

* NOTE: This record contains 77 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 756: contig of 756 bp in length
757 856: gap of 100 bp
857 1580: contig of 724 bp in length
1581 1680: gap of 100 bp
1681 2388: contig of 708 bp in length
2389 2488: gap of 100 bp
2489 3224: contig of 736 bp in length
3225 3324: gap of 100 bp
3325 4064: contig of 740 bp in length
4065 4164: gap of 100 bp
4165 4905: contig of 741 bp in length
4906 5005: gap of 100 bp
5005 5735: contig of 730 bp in length
5736 5835: gap of 100 bp
5836 6590: contig of 755 bp in length
6591 6690: gap of 100 bp
6691 7402: contig of 712 bp in length
7403 7502: gap of 100 bp
7503 8234: contig of 732 bp in length
8235 8334: gap of 100 bp
8335 9055: contig of 722 bp in length
9057 9156: gap of 100 bp
9157 9906: contig of 750 bp in length
9907 10006: gap of 100 bp
10007 10742: contig of 736 bp in length
10743 10842: gap of 100 bp
10843 11586: contig of 744 bp in length
11587 11687: gap of 100 bp
11687 12413: contig of 727 bp in length
12414 12513: gap of 100 bp
12514 13253: contig of 740 bp in length
13254 13353: gap of 100 bp
13354 14100: contig of 747 bp in length
14101 14200: gap of 100 bp
14201 14943: contig of 743 bp in length
14944 15043: gap of 100 bp
15044 15773: contig of 730 bp in length
15774 15873: gap of 100 bp
15874 16589: contig of 716 bp in length
16590 16689: gap of 100 bp
16689 17418: contig of 729 bp in length
17419 17518: gap of 100 bp
17519 18253: contig of 735 bp in length
18254 18354: gap of 100 bp
18354 19092: contig of 739 bp in length
19093 19192: gap of 100 bp
19193 19923: contig of 731 bp in length
19924 20023: gap of 100 bp
20024 20774: contig of 751 bp in length
20775 21608: contig of 734 bp in length
21609 21708: gap of 100 bp
21708 22462: contig of 754 bp in length
22463 22562: gap of 100 bp
22563 23283: contig of 721 bp in length
23284 23383: gap of 100 bp
23384 24120: contig of 737 bp in length
24121 24220: gap of 100 bp
24221 24972: contig of 752 bp in length

24973 25072: gap of 100 bp
25073 25788: contig of 716 bp in length
25789 25888: gap of 100 bp
25889 26634: contig of 746 bp in length
26635 26734: gap of 100 bp
26735 27474: contig of 740 bp in length
27475 27574: gap of 100 bp
27575 28316: contig of 742 bp in length
28317 28416: gap of 100 bp
28417 29131: contig of 715 bp in length
29132 29231: gap of 100 bp
29232 29966: contig of 735 bp in length
29967 30066: gap of 100 bp
30067 30802: contig of 736 bp in length
30803 30902: gap of 100 bp
30903 31633: contig of 731 bp in length
31634 31733: gap of 100 bp
31734 32456: contig of 723 bp in length
32457 32556: gap of 100 bp
32557 33291: contig of 735 bp in length
33292 33391: gap of 100 bp
33392 34117: contig of 726 bp in length
34118 34217: gap of 100 bp
34218 34965: contig of 748 bp in length
34966 35065: gap of 100 bp
35066 35805: contig of 740 bp in length
35806 35905: gap of 100 bp
35906 36649: contig of 744 bp in length
36650 37467: contig of 718 bp in length
37468 37568: gap of 100 bp
37568 38295: contig of 728 bp in length
38296 38395: gap of 100 bp
38396 39123: contig of 728 bp in length
39124 39223: gap of 100 bp
39224 39955: contig of 732 bp in length
39956 40055: gap of 100 bp
40056 40796: contig of 741 bp in length
40797 40896: gap of 100 bp
40897 41629: contig of 733 bp in length
41630 41729: gap of 100 bp
41730 42472: contig of 743 bp in length
42473 42572: gap of 100 bp
42573 43318: contig of 746 bp in length
43319 43418: gap of 100 bp
43419 44166: contig of 748 bp in length
44167 44266: gap of 100 bp
44267 44978: contig of 712 bp in length
44979 45078: gap of 100 bp
45079 45818: contig of 740 bp in length
45819 45918: gap of 100 bp
45919 46636: contig of 718 bp in length
46637 46736: gap of 100 bp
46737 47475: contig of 739 bp in length
47476 47575: gap of 100 bp
47576 48288: contig of 713 bp in length
48289 48388: gap of 100 bp
48389 49105: contig of 717 bp in length
49106 49205: gap of 100 bp
49206 49943: contig of 738 bp in length
49944 50043: gap of 100 bp
50044 50793: contig of 750 bp in length
50794 50893: gap of 100 bp
50894 51657: contig of 764 bp in length
51658 51757: gap of 100 bp
51758 52505: contig of 748 bp in length
52506 52605: gap of 100 bp
52606 53351: contig of 746 bp in length
53352 53451: gap of 100 bp
53452 54214: contig of 763 bp in length
54215 54314: gap of 100 bp
54315 55046: contig of 732 bp in length
55047 55146: gap of 100 bp

```

* 55147 55881: contig of 735 bp in length
* 55882 55981: gap of 100 bp
* 55982 56728: contig of 747 bp in length
* 56729 56828: gap of 100 bp
* 56829 57570: contig of 742 bp in length
* 57571 57670: gap of 100 bp
* 57671 58415: contig of 745 bp in length

Query Match      34.6% Score 18; DB 2; Length 64271;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3 GCAGTCTCTCAGCCTCT 20
Db      43891 GCAGTCTCTCAGCCTCT 43908

RESULT 18
AC018480/c      75162 bp DNA linear HTG 21-DEC-1999
LOCUS DEFINITION Drosophila melanogaster chromosome 2 clone BACR04P15 (D1210)
RPT-98 04.P.15 map 41C-41C strain Y; cn bw sp. *** SEQUENCING IN
PROCESS ***; 60 unordered pieces.
AC018480
AC018480.2 GI:6623916
HTG; HTGS_PHASE1.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 75162)
Celinker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Buenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Katta,K.,
Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacieb,J.M., Park,S., Pfeiffer,B.,
Richards,S., Sethi,H., Svirskaas,R.R., Wan,K.H., Webster,D.,
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
Sequencing of Drosophila melanogaster
Unpublished
2 (bases 1 to 75162)
Celinker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Buenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Katta,K.,
Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacieb,J.M., Park,S., Pfeiffer,B.,
Richards,S., Sethi,H., Svirskaas,R.R., Wan,K.H., Webster,D.,
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
Direct Submission
Submitted (13-DEC-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Dec 21, 1999 this sequence version replaced gi:6563452.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 60 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 509: contig of 509 bp in length
510 589: gap of unknown length
590 1119: contig of 530 bp in length
1120 1199: gap of unknown length
1200 1740: contig of 541 bp in length

```

```

* 38725 41204: contig of 2480 bp in length
* 41205 41284: gap of unknown length
* 41285 44735: contig of 3451 bp in length
* 44736 44815: gap of unknown length
* 44816 45936: contig of 1121 bp in length
* 45937 46016: gap of unknown length
* 46017 47876: contig of 1860 bp in length
* 47877 47956: gap of unknown length
* 47957 50354: contig of 2398 bp in length
* 50355 50434: gap of unknown length
* 50435 54862: contig of 4428 bp in length
* 54863 54942: gap of unknown length
* 54943 58704: contig of 3762 bp in length
* 58705 58784: gap of unknown length
* 58785 62452: contig of 3588 bp in length
* 62453 67667: contig of 5215 bp in length
* 67668 67748: gap of unknown length
* 67749 68309: contig of 562 bp in length
* 68310 68389: gap of unknown length
* 68390 68995: contig of 606 bp in length
* 68996 69075: gap of unknown length
* 69076 69607: contig of 532 bp in length
* 69608 69687: gap of unknown length
* 69688 69906: contig of 219 bp in length
* 69907 69986: gap of unknown length
* 69987 70544: contig of 558 bp in length
* 70545 70624: gap of unknown length
* 70625 71161: contig of 537 bp in length
* 71162 71242: gap of unknown length
* 71243 71818: contig of 577 bp in length
* 71819 71898: gap of unknown length
* 71899 72368: contig of 470 bp in length
* 72369 72448: gap of unknown length
* 72449 73062: contig of 614 bp in length
* 73063 73142: gap of unknown length
* 73143 73763: contig of 621 bp in length
* 73764 73843: gap of unknown length
* 73844 74543: contig of 700 bp in length
* 74544 74624: gap of unknown length
* 74625 75162: contig of 539 bp in length.

FEATURES
SOURCE
1..75162
location/Qualifiers
    /organism="Drosophila melanogaster"
    /mol_type="genomic DNA"
    /strain="y: cn bw sp"
    /db_xref="taxon:7227"
    /chromosome="2"
    /map="41C-41C"
    /clone="BACR04P15 (D1210) RPCI-98 04.P.15"
    /clone_11b="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACe3.6"

BASE COUNT      19287 a 15828 c 15319 g 19977 t 4751 others
ORIGIN
Query Match      34.6%; Score 18; DB 2; Length 75162;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      21 CATGCTCTCGTCTGCTGC 38
Db      65558 CATGCTCTCGTCTGCTGC 65541

RESULT 19
AC010568/c
LOCUS   AC010568      99277 bp   DNA       linear   HTG 29-MAY-2000
DEFINITION
Drosophila melanogaster clone RPCI98-1716, *** SEQUENCING IN
PROGRAMS ***. 43 unordered pieces.
ACCESSION
AC010568
VERSION
AC010568.5 GI:8101320
KEYWORDS
HTG; HTGS_PHASE1.

```

```

SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Braconyera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 99277)
REFERENCE
AUTHORS
Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,
Bodda,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Doman-Rashid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferriguo,D.,
Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
Guevara,M., Harris,K., Hernandez,J., Hodgson,A., Hogue,M.,
Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
Lichter,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
Shah,B., Shen,H., Simon,M., Sparks,A., Stamps,A., Sugang,R.,
Taber,P., Taylor,T., Vasquez,L., Vinson,R., Vo,O., Wabba,M.,
Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
Worley,K., Wren,J., Wrenstock,G., Yu,W., Zhou,X., Nelson,D. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 99277)
Worley,K.C.
Submitted (16-SEP-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 29, 2000 this sequence version replaced gi:696776.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc.help@bcm.tmc.edu
----- Project Information
Center project name: DRFQ
Center clone name: RPCI98-1716
----- Summary Statistics
Sequencing vector: M13; 108821
Chemistry: Dye-primer Bodypy; 14% of reads
Chemistry: Dye-terminator Big Dye; 4% of reads
Assembly program: Phrap; version 0.980611
Consensus quality: 69137 bases at least Q40
Consensus quality: 82549 bases at least Q30
Consensus quality: 89824 bases at least Q20
Estimated insert size: 95034; sum-of-contigs estimation
Quality coverage: 2.8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 43 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
10767: contig of 10767 bp in length
10768 10867: gap of unknown length
10868 16333: contig of 5466 bp in length
16334 16433: gap of unknown length
28333 28333: contig of 11900 bp in length
28334 28433: gap of unknown length
28434 30669: contig of 2236 bp in length
30670 30769: gap of unknown length
30770 33570: contig of 2801 bp in length
33571 33671: gap of unknown length
33672 36538: contig of 2868 bp in length
36539 36638: gap of unknown length

```

```

* 3639 38818: contig of 2180 bp in length
* 38819 38918: gap of unknown length
* 40501 40501: contig of 1583 bp in length
* 40602 40601: gap of unknown length
* 42385 42385: contig of 1784 bp in length
* 42485 42485: gap of unknown length
* 45266 45266: contig of 3041 bp in length
* 45527 45527: gap of unknown length
* 47755 47755: contig of 2129 bp in length
* 47855 47855: gap of unknown length
* 49531 49531: contig of 1675 bp in length
* 49631 49630: gap of unknown length
* 51574 51574: contig of 1944 bp in length
* 51674 51674: gap of unknown length
* 53258 53258: contig of 1584 bp in length
* 53359 53358: gap of unknown length
* 55214 55214: contig of 1856 bp in length
* 55315 55314: gap of unknown length
* 58025 58025: contig of 2711 bp in length
* 58125 58125: gap of unknown length
* 58126 58126: contig of 2056 bp in length
* 60182 60281: gap of unknown length
* 60282 62276: contig of 1995 bp in length
* 62377 62376: gap of unknown length
* 63994 63994: contig of 1618 bp in length
* 63995 64094: gap of unknown length
* 64095 66601: contig of 2507 bp in length
* 66602 66701: gap of unknown length
* 66702 67904: contig of 1203 bp in length
* 67905 68004: gap of unknown length
* 68005 69185: contig of 1181 bp in length
* 69186 69285: gap of unknown length
* 69286 70508: contig of 1223 bp in length
* 70509 70608: gap of unknown length
* 70609 72036: contig of 1428 bp in length
* 72037 72136: gap of unknown length
* 72137 74181: contig of 2045 bp in length
* 74182 74281: gap of unknown length
* 74282 75295: contig of 1014 bp in length
* 75296 75395: gap of unknown length
* 75396 76878: contig of 1483 bp in length
* 76879 76978: gap of unknown length
* 76979 78393: contig of 1415 bp in length
* 78394 78493: gap of unknown length
* 78494 80210: contig of 1717 bp in length
* 80211 80310: gap of unknown length
* 80311 81493: contig of 1183 bp in length
* 81494 81593: gap of unknown length
* 81594 82783: contig of 1190 bp in length
* 82784 82883: gap of unknown length
* 82884 84809: contig of 1926 bp in length
* 84810 84909: gap of unknown length
* 84910 86592: contig of 1683 bp in length
* 86593 86692: gap of unknown length
* 86693 87752: contig of 1060 bp in length
* 87753 87852: gap of unknown length
* 87853 89155: contig of 1303 bp in length
* 89156 89255: gap of unknown length
* 89256 90443: contig of 1188 bp in length
* 90444 90543: gap of unknown length
* 90544 91893: contig of 1350 bp in length
* 91894 91993: gap of unknown length
* 91994 93070: contig of 1077 bp in length
* 93071 93170: gap of unknown length
* 93171 94175: contig of 1005 bp in length
* 94176 94275: gap of unknown length
* 94276 95483: contig of 1208 bp in length
* 95484 95583: gap of unknown length
* 95584 96799: contig of 1216 bp in length
* 96799 96898: gap of unknown length
* 96899 98039: contig of 1200 bp in length
* 98100 98199: gap of unknown length
* 99277: contig of 1078 bp in length.

```

```

FEATURES
    source
        Location/Qualifiers
            1..99277
                /organism="Drosophila melanogaster"
                /mol_type="genomic DNA"
                /db_xref="taxon:7227"
                /clone="RPC198-1716"
    misc_feature
        complement(26106..26205)
        /note="assembly:fragment:contig35; vector:Vector_pBACE3.6;
        vector_side:left"
    misc_feature
        41196..41295
        /note="assembly:fragment:contig36; vector:Vector_pBACE3.6;
        vector_side:right"
BASE COUNT      28182 a 19521 c 18383 g 28965 t 4226 others
ORIGIN
Query Match      34.6%; Score 18; DB 2; Length 99277;
Best Local Similarity 100.0%; Pred No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      21 CATCGCTCTGTCGCTGC 38
Db      35859 CATCGCTCTGTCGCTGC 35842

RESULT 20
AC008227      101579 bp      DNA      linear      HTG 20-SEP-1999
LOCUS      Drosophila melanogaster chromosome 3 clone BACR05008 (D822) RCI-98
DEFINITION      05.0.8 map 84p-84p strain y: cn bw sp, *** SEQUENCING IN PROGRESS
                *** 56 unordered pieces.
ACCESSION      AC008227
VERSION      AC008227.2 GI:5670434
KEYWORDS      HTG; HTGS PHASE1.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM      Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 101579)
AUTHORS      Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champagne,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummachi,S.R., Kairo,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomocan,M.A., Mazda,P.,
Moshirefi,A.R., Moshirefi,M., Nixon,K., Paclebo,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequerra,A., Sethi,H., Snir,E.,
Svitskaas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
TITLE      Sequencing of Drosophila melanogaster
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 101579)
AUTHORS      Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champagne,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummachi,S.R., Kairo,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomocan,M.A., Mazda,P.,
Moshirefi,A.R., Moshirefi,M., Nixon,K., Paclebo,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequerra,A., Sethi,H., Snir,E.,
Svitskaas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
TITLE      Direct Subinjection
JOURNAL      Submitted (29-JUL-1999) Drosophila Genome Center, Lawrence Berkeley
COMMENT      Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Aug 2, 1999 this sequence version replaced gi:5649193.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this subinjection meet
the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

```

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
589: contig of 589 bp in length
590: gap of unknown length
670: contig of 521 bp in length
1190: contig of 521 bp in length
1270: gap of unknown length
1271: contig of 538 bp in length
1808: gap of unknown length
1889: contig of 595 bp in length
2484: gap of unknown length
2563: gap of unknown length
3225: contig of 662 bp in length
3305: gap of unknown length
3326: contig of 523 bp in length
3828: gap of unknown length
3908: contig of 527 bp in length
4435: gap of unknown length
4515: contig of 639 bp in length
5154: gap of unknown length
5234: contig of 583 bp in length
5817: gap of unknown length
5897: contig of 1086 bp in length
5898: gap of unknown length
6983: contig of 952 bp in length
7063: gap of unknown length
8015: contig of 952 bp in length
8095: gap of unknown length
8096: contig of 689 bp in length
8784: gap of unknown length
8864: contig of 659 bp in length
9523: gap of unknown length
9603: contig of 934 bp in length
10537: gap of unknown length
10617: contig of 812 bp in length
11429: gap of unknown length
11509: contig of 700 bp in length
12209: gap of unknown length
12210: contig of 649 bp in length
12939: gap of unknown length
13018: contig of 1183 bp in length
14201: gap of unknown length
14281: contig of 833 bp in length
15114: gap of unknown length
15195: contig of 1508 bp in length
16702: gap of unknown length
16783: contig of 825 bp in length
17607: gap of unknown length
17688: contig of 1322 bp in length
19009: gap of unknown length
19089: contig of 1992 bp in length
21081: gap of unknown length
21161: contig of 1312 bp in length
22473: gap of unknown length
22553: contig of 1272 bp in length
23825: gap of unknown length
23906: contig of 1378 bp in length
25283: gap of unknown length
25363: contig of 1454 bp in length
26817: gap of unknown length
26897: gap of unknown length
28915: contig of 2018 bp in length
28995: gap of unknown length
30314: contig of 1319 bp in length
30394: gap of unknown length
31753: contig of 1359 bp in length
31754: gap of unknown length
31834: contig of 1241 bp in length
33075: gap of unknown length
33155: contig of 2172 bp in length
35326: gap of unknown length
35406: contig of 2167 bp in length
37573: gap of unknown length
37654: contig of 2099 bp in length
39753: gap of unknown length
41517: contig of 1685 bp in length

```

```

* 41518 41597: gap of unknown length
* 41598 44746: contig of 3149 bp in length
* 44747 44826: gap of unknown length
* 44827 48239: contig of 3413 bp in length
* 48240 48319: gap of unknown length
* 48320 52203: contig of 3884 bp in length
* 52204 52283: gap of unknown length
* 52284 59738: contig of 7455 bp in length
* 59739 59818: gap of unknown length
* 59819 67723: contig of 7905 bp in length
* 67724 67803: gap of unknown length
* 67804 72670: contig of 4867 bp in length
* 72671 72750: gap of unknown length
* 72751 91537: contig of 18787 bp in length
* 91538 91617: gap of unknown length
* 91618 92277: contig of 660 bp in length
* 92278 92357: gap of unknown length
* 92358 93027: contig of 670 bp in length
* 93028 93107: gap of unknown length
* 93108 93733: contig of 625 bp in length
* 93733 93812: gap of unknown length
* 93813 94380: contig of 568 bp in length
* 94381 94460: gap of unknown length
* 94461 95199: contig of 739 bp in length
* 95200 95279: gap of unknown length
* 95280 95663: contig of 384 bp in length
* 95664 95743: gap of unknown length
* 95744 96388: contig of 645 bp in length
* 96389 96468: gap of unknown length
* 96469 97219: contig of 751 bp in length
* 97220 97299: gap of unknown length
* 97300 98072: contig of 773 bp in length
* 98073 98152: gap of unknown length
* 98153 98805: contig of 653 bp in length
* 98806 98885: gap of unknown length
* 98886 99435: contig of 550 bp in length
* 99436 99515: gap of unknown length
* 100108 100107: contig of 592 bp in length
* 100109 100187: gap of unknown length
* 100188 100895: contig of 708 bp in length
* 100896 100975: gap of unknown length
* 100976 100976: contig of 604 bp in length.
* Location/Qualifiers
* 1. 101579
* /organism="Drosophila melanogaster"
* /mol_type="genomic DNA"
* /strain="Y; cn Bw bp"
* /db_xref="taxon:7227"
* /chromosome="3"
* /map="84D-84D"
* /clone="BACR05008 (D822) RPCT-98.05.O.8"
* /clone_id="RPCT-98 (Roswell Park Cancer Institute
* Drosophila melanogaster BAC library, partial EcoRI in
* pBACe3.6"
*
BASE COUNT 29377 a 20734 c 19813 g 27246 t 4409 others
ORIGIN
Query Match 34.6%; Score 18; DB 2; Length 101579;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 21 CATGCTCTGTCGCTGC 38
Db 67542 CATGCTCTGTCGCTGC 67559
RESULT 21
AC098797 117714 bp DNA linear INV 01-JAN-2003
LOCUS AC098797
DEFINITION Leishmania major chromosome 35 clone LB01304 strain Friedlin,
complete sequence.
ACCESSION AC098797
VERSION AC098797.3 GI:18139378

```


KEYWORDS HTG
SOURCE Leishmania major
ORGANISM Leishmania major
REFERENCE 1 (bases 1 to 117714)
AUTHORS Myler,P.J., Sisk,E., Cawthra,J., Handley,F., Vogt,C., Robertson,L., McDonagh,P., Ivens,A., Nguyen,D., Munden,H. and Stuart,K.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-2001) Seattle Biomedical Research Institution, 4
REFERENCE 2 (bases 1 to 117714)
AUTHORS Myler,P.J., Sisk,E., Nelson,S., Mack,J., Vogt,C., Robertson,L., Worthey,L., Ivens,A., Marty,A., Munden,H., Stuart,K., Rinta,J. and Seyler,A.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-2001) Seattle Biomedical Research Institution, 4
REFERENCE 3 (bases 1 to 117714)
AUTHORS Myler,P.J., Sisk,E., Nelson,S., Mack,J., Vogt,C., Robertson,L., Worthey,L., Ivens,A., Marty,A., Munden,H., Stuart,K., Rinta,J. and Seyler,A.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-2002) Seattle Biomedical Research Institution, 4
REFERENCE 4 (bases 1 to 117714)
AUTHORS Myler,P.J., Sisk,E., Nelson,S., Mack,J., Vogt,C., Robertson,L., Worthey,L., Ivens,A., Marty,A., Munden,H., Stuart,K., Rinta,J. and Seyler,A.
TITLE Direct Submission
JOURNAL Submitted (01-JAN-2003) Seattle Biomedical Research Institution, 4
COMMENT Nickerson Street, Seattle, WA 98109-1651, USA
FEATURES
source Location/Qualifiers
1..117714
/organism="Leishmania major"
/mol_type="genomic DNA"
/strain="Friedlin"
/db_xref="taxon:5664"
/chromosome="J5"
/clone="LB01304"
BASE COUNT 22292 a 34596 c 33718 g 27108 t
ORIGIN
Query Match 34.6% Score 18; DB 3; Length 117714;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 20 TCATCGCTCTCGTCGCTG 37
Db 98830 TCATCGCTCTCGTCGCTG 98847
RESULT 22
LOCUS AC009746
DEFINITION Drosophila melanogaster chromosome 2 clone BACR03D04 (D1041)
AC009746 123010 bp DNA linear HTG 20-MAR-2000
RPTC-98.03.D.4 map 40D-40D strain Y; cn bw sp, *** SEQUENCING IN
PROGRES ***; 80 unordered pieces.
AC009746
AC009746.10 GI:7264757
HTG; HTGS PHASE1.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydriidae; Drosophilidae; Drosophila.
1 (bases 1 to 123010)
Celiker,S.E., Abayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Buenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,
Richards,S., Sethi,H., Svirskaas,R.R., Wan,K.H., Webster,D.,
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
Sequencing of Drosophila melanogaster
Unpublished
2 (bases 1 to 123010)
Celiker,S.E., Abayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Buenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomocan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequerra,A., Sethi,H., Shtir,E.,
Svirskaas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT For further information about this sequence replaced gi:7105464.
On Mar 20, 2000 this sequence version replaced gi:7105464.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive web site (<http://www.fruitfly.org/sequence/>) or send email
to bdgp@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 80 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
497 496: contig of 496 bp in length
577 576: gap of unknown length
975 974: contig of 398 bp in length
1055 1054: gap of unknown length
1417 1416: contig of 362 bp in length
1497 1496: gap of unknown length
2155 2154: contig of 658 bp in length
2235 2234: gap of unknown length
2977 2976: contig of 742 bp in length
3057 3056: gap of unknown length
3885 3884: contig of 828 bp in length
3965 3964: gap of unknown length
4689 4688: contig of 724 bp in length
4769 4768: gap of unknown length
8376 8375: contig of 3607 bp in length
8456 8455: gap of unknown length
11189 11188: contig of 2734 bp in length
11269 11268: gap of unknown length
11270 11269: contig of 3794 bp in length
15063 15062: gap of unknown length
15143 15142: gap of unknown length
15144 15143: contig of 3450 bp in length
18593 18592: gap of unknown length
18673 18672: gap of 3552 bp in length
22266 22265: gap of unknown length
22346 22345: gap of unknown length
22346 22345: contig of 5105 bp in length
27451 27450: gap of unknown length
33693 33692: contig of 6153 bp in length
33694 33693: gap of unknown length
33773 33772: gap of unknown length
33774 33773: contig of 6737 bp in length
40511 40510: gap of unknown length
40591 40590: gap of unknown length
48557 48556: contig of 7967 bp in length
48558 48557: gap of unknown length
48637 48636: gap of unknown length
60190 60189: contig of 11553 bp in length
60270 60269: gap of unknown length
60191 60190: gap of unknown length
70860 70859: contig of 10510 bp in length
70861 70860: gap of unknown length
82464 82463: contig of 11604 bp in length
82465 82464: gap of unknown length
83243 83242: contig of 719 bp in length
83264 83263: gap of unknown length
83343 83342: gap of unknown length
83944 83943: contig of 601 bp in length

KEYWORDS HTG
SOURCE Leishmania major
ORGANISM Leishmania major
REFERENCE 1 (bases 1 to 117714)
AUTHORS Myler,P.J., Sisk,E., Cawthra,J., Handley,F., Vogt,C., Robertson,L., McDonagh,P., Ivens,A., Nguyen,D., Munden,H. and Stuart,K.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-2001) Seattle Biomedical Research Institution, 4
REFERENCE 2 (bases 1 to 117714)
AUTHORS Myler,P.J., Sisk,E., Nelson,S., Mack,J., Vogt,C., Robertson,L., Worthey,L., Ivens,A., Marty,A., Munden,H., Stuart,K., Rinta,J. and Seyler,A.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-2001) Seattle Biomedical Research Institution, 4
REFERENCE 3 (bases 1 to 117714)
AUTHORS Myler,P.J., Sisk,E., Nelson,S., Mack,J., Vogt,C., Robertson,L., Worthey,L., Ivens,A., Marty,A., Munden,H., Stuart,K., Rinta,J. and Seyler,A.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-2002) Seattle Biomedical Research Institution, 4
REFERENCE 4 (bases 1 to 117714)
AUTHORS Myler,P.J., Sisk,E., Nelson,S., Mack,J., Vogt,C., Robertson,L., Worthey,L., Ivens,A., Marty,A., Munden,H., Stuart,K., Rinta,J. and Seyler,A.
TITLE Direct Submission
JOURNAL Submitted (01-JAN-2003) Seattle Biomedical Research Institution, 4
COMMENT Nickerson Street, Seattle, WA 98109-1651, USA
FEATURES
source Location/Qualifiers
1..117714
/organism="Leishmania major"
/mol_type="genomic DNA"
/strain="Friedlin"
/db_xref="taxon:5664"
/chromosome="J5"
/clone="LB01304"
BASE COUNT 22292 a 34596 c 33718 g 27108 t
ORIGIN
Query Match 34.6% Score 18; DB 3; Length 117714;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 20 TCATCGCTCTCGTCGCTG 37
Db 98830 TCATCGCTCTCGTCGCTG 98847
RESULT 22
LOCUS AC009746
DEFINITION Drosophila melanogaster chromosome 2 clone BACR03D04 (D1041)
AC009746 123010 bp DNA linear HTG 20-MAR-2000
RPTC-98.03.D.4 map 40D-40D strain Y; cn bw sp, *** SEQUENCING IN
PROGRES ***; 80 unordered pieces.
AC009746
AC009746.10 GI:7264757
HTG; HTGS PHASE1.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydriidae; Drosophilidae; Drosophila.
1 (bases 1 to 123010)
Celiker,S.E., Abayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Buenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,
Richards,S., Sethi,H., Svirskaas,R.R., Wan,K.H., Webster,D.,
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
Sequencing of Drosophila melanogaster
Unpublished
2 (bases 1 to 123010)
Celiker,S.E., Abayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Buenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomocan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequerra,A., Sethi,H., Shtir,E.,
Svirskaas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT For further information about this sequence replaced gi:7105464.
On Mar 20, 2000 this sequence version replaced gi:7105464.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive web site (<http://www.fruitfly.org/sequence/>) or send email
to bdgp@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 80 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
497 496: contig of 496 bp in length
577 576: gap of unknown length
975 974: contig of 398 bp in length
1055 1054: gap of unknown length
1417 1416: contig of 362 bp in length
1497 1496: gap of unknown length
2155 2154: contig of 658 bp in length
2235 2234: gap of unknown length
2977 2976: contig of 742 bp in length
3057 3056: gap of unknown length
3885 3884: contig of 828 bp in length
3965 3964: gap of unknown length
4689 4688: contig of 724 bp in length
4769 4768: gap of unknown length
8376 8375: contig of 3607 bp in length
8456 8455: gap of unknown length
11189 11188: contig of 2734 bp in length
11269 11268: gap of unknown length
11270 11269: contig of 3794 bp in length
15063 15062: gap of unknown length
15143 15142: gap of unknown length
15144 15143: contig of 3450 bp in length
18593 18592: gap of unknown length
18673 18672: gap of 3552 bp in length
22266 22265: gap of unknown length
22346 22345: gap of unknown length
22346 22345: contig of 5105 bp in length
27451 27450: gap of unknown length
33693 33692: contig of 6153 bp in length
33694 33693: gap of unknown length
33773 33772: gap of unknown length
33774 33773: contig of 6737 bp in length
40511 40510: gap of unknown length
40591 40590: gap of unknown length
48557 48556: contig of 7967 bp in length
48558 48557: gap of unknown length
48637 48636: gap of unknown length
60190 60189: contig of 11553 bp in length
60270 60269: gap of unknown length
60191 60190: gap of unknown length
70860 70859: contig of 10510 bp in length
70861 70860: gap of unknown length
82464 82463: contig of 11604 bp in length
82465 82464: gap of unknown length
83243 83242: contig of 719 bp in length
83264 83263: gap of unknown length
83343 83342: gap of unknown length
83944 83943: contig of 601 bp in length


```

* 83945 84024: gap of unknown length
* 84025 84537: contig of 513 bp in length
* 84538 84617: gap of unknown length
* 84618 84853: contig of 236 bp in length
* 84854 84933: gap of unknown length
* 84934 85721: contig of 788 bp in length
* 85722 85801: gap of unknown length
* 85802 86606: contig of 805 bp in length
* 86607 86686: gap of unknown length
* 86687 87380: contig of 694 bp in length
* 87381 87460: gap of unknown length
* 87461 87953: contig of 493 bp in length
* 87954 88033: gap of unknown length
* 88034 88563: contig of 530 bp in length
* 88564 88643: gap of unknown length
* 88644 89377: contig of 734 bp in length
* 89378 89457: gap of unknown length
* 89458 90050: contig of 593 bp in length
* 90051 90130: gap of unknown length
* 90131 90973: contig of 843 bp in length
* 90974 91053: gap of unknown length
* 91054 91730: contig of 677 bp in length
* 91731 91810: gap of unknown length
* 91811 92843: contig of 1033 bp in length
* 92844 92923: gap of unknown length
* 92924 93540: contig of 617 bp in length
* 93541 93620: gap of unknown length
* 93621 94200: contig of 580 bp in length
* 94201 94280: gap of unknown length
* 94281 94644: contig of 364 bp in length
* 94645 94724: gap of unknown length
* 94725 95220: contig of 496 bp in length
* 95221 95300: gap of unknown length
* 95301 95654: contig of 354 bp in length
* 95655 95734: gap of unknown length
* 95735 96325: contig of 591 bp in length
* 96326 96405: gap of unknown length
* 96406 97058: contig of 653 bp in length
* 97059 97138: gap of unknown length
* 97139 97533: contig of 395 bp in length
* 97534 97613: gap of unknown length
* 97614 97970: contig of 357 bp in length
* 97971 98050: gap of unknown length
* 98051 98757: contig of 707 bp in length
* 98758 98837: gap of unknown length
* 98838 99579: contig of 742 bp in length
* 99580 99659: gap of unknown length
* 99660 100277: contig of 618 bp in length
* 100278 100357: gap of unknown length
* 100358 101128: contig of 771 bp in length
* 101129 101208: gap of unknown length
* 101209 102301: contig of 1093 bp in length
* 102302 102381: gap of unknown length
* 102382 102818: contig of 437 bp in length
* 102819 102898: gap of unknown length
* 102899 103305: contig of 407 bp in length
* 103306 103385: gap of unknown length
* 103386 103989: contig of 604 bp in length
* 103990 104069: gap of unknown length
* 104070 104735: contig of 666 bp in length
* 104736 104815: gap of unknown length
* 104816 105416: contig of 601 bp in length
* 105417 105496: gap of unknown length
* 105497 105886: contig of 390 bp in length
* 105887 105966: gap of unknown length
* 105967 106460: contig of 494 bp in length
* 106461 106540: gap of unknown length
* 106541 107486: contig of 946 bp in length
* 107487 107566: gap of unknown length
* 107567 107911: contig of 345 bp in length
* 107912 107991: gap of unknown length
* 107992 108893: contig of 902 bp in length
* 108894 108973: gap of unknown length

```

```

* 108974 109636: contig of 663 bp in length
* 109637 109716: gap of unknown length
* 109717 110025: contig of 309 bp in length
* 110026 110105: gap of unknown length
* 110106 110514: contig of 409 bp in length
* 110515 110594: gap of unknown length
* 110595 110842: contig of 248 bp in length
* 110843 110922: gap of unknown length
* 110923 111715: contig of 793 bp in length
* 111716 111795: gap of unknown length
* 111796 112290: contig of 495 bp in length
* 112291 112370: gap of unknown length
* 112371 113102: contig of 732 bp in length
* 113103 113182: gap of unknown length
* 113183 113446: contig of 264 bp in length
* 113447 113526: gap of unknown length
* 113527 114149: contig of 623 bp in length
* 114150 114229: gap of unknown length
* 114230 114773: contig of 544 bp in length
* 114774 114853: gap of unknown length
* 114854 115250: contig of 397 bp in length
* 115251 115330: gap of unknown length
* 115331 116104: contig of 774 bp in length
* 116105 116184: gap of unknown length
* 116185 116562: contig of 378 bp in length
* 116563 116642: gap of unknown length
* 116643 116857: contig of 215 bp in length
* 116858 116937: gap of unknown length
* 116938 117574: contig of 637 bp in length
* 117575 117654: gap of unknown length
* 117655 118239: contig of 585 bp in length
* 118240 118319: gap of unknown length
* 118320 118629: contig of 310 bp in length
* 118630 118709: gap of unknown length
* 118710 119196: contig of 487 bp in length
* 119197 119276: gap of unknown length

```

```

Query Match      34.6%  Score 18:  DB 2:  Length 123010;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      21  CATCGCTCTCGTCGCTGC 38
Db      30732 CATCGCTCTCGTCGCTGC 30749

```

```

RESULT 23
AC105442/c      124612 bp      DNA      linear      HTG 29-JAN-2002
LOCUS
DEFINITION
Leishmania major chromosome 35 clone LB00329 strain Friedlin, ***
SEQUENCING IN PROGRESS ***, 3 ordered pieces.

```

```

AC105442
AC105442.2  GI:18390199
VERSION
HTG; HTGS PHASE2.
SOURCE
Leishmania major
ORGANISM
Eukaryota; Elenozoa; Kinetoplastida; Trypanosomatidae;
Leishmania.

```

```

REFERENCE
1 (bases 1 to 124612)
Myler,P.J., Sisk,E., Mack,U., Nelson,S., Vogt,C., Robertson,L.,
Wortheley,L., Ivens,A., Marty,A., Munden,H., Stuart,K., Seyler,A. and
Rine,J.

```

```

TITLE
Direct Submision
JOURNAL
Submitted (05-JAN-2002) Seattle Biomedical Research Institution, 4
Nickerson Street, Seattle, WA 98109-1651, USA
On Jan 29, 2002 this sequence version replaced gi:18071338.

```

```

COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submittor.
* This sequence will be replaced

```

* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 103780: contig of 103780 bp in length
* 103781 103930: gap of unknown length
* 103931 106086: contig of 2156 bp in length
* 106087 106236: gap of unknown length
* 106237 124612: contig of 18376 bp in length.
Location/Qualifiers
source
1. 124612
/organism="Leishmania major"
/mol_type="genomic DNA"
/strain="Friedlin"
/db_xref="taxon:5664"
/chromosome="35"
/clone="LB00329"
BASE COUNT 29211 a 35011 c 36514 g 23576 t 300 others
ORIGIN
Query Match 34.6%; Score 18; DB 2; Length 124612;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 20 TCATCGCTCTCGTCGCTG 37
Db 105161 TCATCGCTCTCGTCGCTG 105144
RESULT 24 127091 bp DNA linear HTG 28-FEB-2000
AC009843/c Drosophila melanogaster chromosome 3 clone BACR01G11 (D1045)
LOCUS RPL1-98 01.G.11 map 80B1-80D2 strain Y; cn bw sp. *** SEQUENCING IN
DEFINITION PROGRESS ***; 93 unordered pieces.
AC009843
AC009843.7 GI:7107697
HTG: HTGS PHASE1.
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 127091)
Celiker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Lee,B., Lewis,S., Li,P., Lomocan,W.A., Mazda,P.,
Moshrefi,M., Nixon,K., Pacleib,J.M., Park,S., Pfeiffer,B.,
Richards,S., Sethi,H., Svirekas,R.R., Wan,K.H., Webster,D.,
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
Sequencing of Drosophila melanogaster
Unpublished
2 (bases 1 to 127091)
Celiker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomocan,W.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleib,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequerra,A., Sethi,H., Snir,E.,
Svirekas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
Direct Submission
Submitted (03-SEP-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Feb 28, 2000 this sequence version replaced g1:6980150.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 93 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 562: contig of 562 bp in length
563 642: gap of unknown length
643 1276: contig of 634 bp in length
1277 1356: gap of unknown length
1357 1983: contig of 627 bp in length
1984 2063: gap of unknown length
2064 2313: contig of 250 bp in length
2314 2393: gap of unknown length
2394 2865: contig of 472 bp in length
2866 2945: gap of unknown length
2946 3493: contig of 548 bp in length
3494 3573: gap of unknown length
3574 4071: contig of 498 bp in length
4072 4151: gap of unknown length
4152 4897: contig of 746 bp in length
4898 4977: gap of unknown length
4978 5419: contig of 442 bp in length
5419 5499: gap of unknown length
5499 5500: gap of 972 bp in length
5500 6471: contig of 972 bp in length
6472 6551: gap of unknown length
6551 7505: contig of 954 bp in length
7506 7585: gap of unknown length
7586 8204: contig of 619 bp in length
8205 8284: gap of unknown length
8285 9209: contig of 925 bp in length
9210 9289: gap of unknown length
9290 9981: contig of 692 bp in length
9982 10061: gap of unknown length
10061 11280: contig of 1139 bp in length
11280 11280: gap of unknown length
11281 11987: contig of 707 bp in length
11987 12067: gap of unknown length
12068 12786: contig of 719 bp in length
12787 12866: gap of unknown length
12867 13456: contig of 590 bp in length
13457 13537: gap of unknown length
13537 14333: contig of 797 bp in length
14334 14413: gap of unknown length
14413 15065: contig of 652 bp in length
15066 15145: gap of unknown length
15145 15574: contig of 429 bp in length
15574 15654: gap of unknown length
15655 16170: contig of 516 bp in length
16170 16250: gap of unknown length
16250 16868: contig of 618 bp in length
16868 16948: gap of unknown length
16948 17712: contig of 764 bp in length
17712 17793: gap of unknown length
17793 19275: gap of 1483 bp in length
19275 19355: gap of unknown length
19356 20073: contig of 718 bp in length
20074 20153: gap of unknown length
20154 20797: contig of 644 bp in length
20798 20877: gap of unknown length
20877 22421: contig of 1544 bp in length
22421 22501: gap of unknown length
22501 23477: contig of 976 bp in length
23477 23558: gap of unknown length
23558 24589: contig of 1022 bp in length
24589 24669: gap of unknown length
24670 25381: contig of 712 bp in length
25382 25451: gap of unknown length
25451 26619: contig of 1158 bp in length
26619 26620: gap of unknown length
26620 27921: contig of 1222 bp in length
27921 28001: gap of unknown length
28001 28959: contig of 958 bp in length

```

* 28960 29039: gap of unknown length
* 29040 30564: contig of 1525 bp in length
* 30565 30644: gap of unknown length
* 30645 31644: contig of 1000 bp in length
* 31645 31724: gap of unknown length
* 31725 33109: contig of 1385 bp in length
* 33110 33189: gap of unknown length
* 33190 34898: contig of 1709 bp in length
* 34899 34979: gap of unknown length
* 34979 35797: contig of 819 bp in length
* 35798 35877: gap of unknown length
* 35878 37188: contig of 1311 bp in length
* 37189 37268: gap of unknown length
* 37269 38152: contig of 884 bp in length
* 38153 38232: gap of unknown length
* 38233 40088: contig of 1856 bp in length
* 40089 40168: gap of unknown length
* 40169 41888: contig of 1720 bp in length
* 41889 41968: gap of unknown length
* 41969 43562: contig of 1594 bp in length
* 43563 43642: gap of unknown length
* 43643 44695: contig of 1053 bp in length
* 44696 44775: gap of unknown length
* 44776 46594: contig of 1819 bp in length
* 46595 46674: gap of unknown length
* 46675 48005: contig of 1331 bp in length
* 48006 48085: gap of unknown length
* 48086 49697: contig of 1612 bp in length
* 49698 49777: gap of unknown length
* 49778 50928: contig of 1151 bp in length
* 50929 51008: gap of unknown length
* 51009 53046: contig of 2038 bp in length
* 53047 53126: gap of unknown length
* 53127 54856: contig of 1730 bp in length
* 54857 54936: gap of unknown length
* 54937 56265: contig of 1329 bp in length
* 56266 56345: gap of unknown length
* 56346 57715: contig of 1370 bp in length
* 57716 57795: gap of unknown length
* 57796 59529: contig of 1734 bp in length
* 59530 59609: gap of unknown length
* 59610 61234: contig of 1625 bp in length
* 61235 61314: gap of unknown length
* 61315 63824: contig of 2510 bp in length
* 63825 63904: gap of unknown length
* 63905 6453: contig of 2549 bp in length
* 6454 66533: gap of unknown length
* 66534 71049: contig of 4516 bp in length
* 71050 71129: gap of unknown length
* 71130 75699: contig of 4570 bp in length
* 75700 75779: gap of unknown length
* 75780 79646: contig of 3867 bp in length
* 79647 79726: gap of unknown length
* 79727 84818: contig of 5092 bp in length
* 84819 84898: gap of unknown length
* 84899 90429: contig of 5531 bp in length
* 90430 90509: gap of unknown length
* 90510 93957: contig of 3448 bp in length
* 93958 94037: gap of unknown length
* 94038 100461: contig of 6424 bp in length
* 100462 100541: gap of unknown length
* 100542 106233: contig of 5692 bp in length
* 106234 106313: gap of unknown length
* 106314 106987: contig of 674 bp in length
* 106988 107067: gap of unknown length
* 107068 107780: contig of 713 bp in length
* 107781 107860: gap of unknown length
* 107861 108540: contig of 680 bp in length
* 108541 108620: gap of unknown length
* 108621 109185: contig of 565 bp in length
* 109186 109265: gap of unknown length
* 109266 109901: contig of 636 bp in length
* 109902 109981: gap of unknown length

```

```

Query Match      34.6% Score 18; DB 2; Length 127091;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      21 CATCGCTCTCGTCTGC 38
DB      88489 CATCGCTCTCGTCTGC 88472

RESULT 25
AC010915
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 137841)
Celiker,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Banton,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Chape,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dreenek,D., Fartan,D.,
Fiertera,S., Friese,E., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houch,T., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegwam,C., Jalili,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
McIntosh,C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
Phonanaong,S., Piltman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome 4, region 101F-102F
Unpublished (1998)
2 (bases 1 to 137841)
Celiker,S.E., Abayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Buenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Fartan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomtan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Shit,E.,
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
Direct Submission
Submitted (27-SEP-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Jan 10, 2002 this sequence version replaced gi:7116216.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdgp@fruitfly.berkeley.edu.

```

FEATURES
source
1. 137841
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="Y: cn bw sp"
/db_xref="taxon:7227"
/chromosome="4"
/map="101F-102F"
/clone="BACR05L22 (D1068)"
/clone_1b="RPC1-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBAC3.6)"

BASE COUNT
44191 a 25673 c 24261 g 43716 t
ORIGIN

Query Match 34.6%; Score 18; DB 3; Length 137841;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 CATCCCTCTGCTGCTGC 38
|||||
Db 12415 CATCCCTCTGCTGCTGC 12432

RESULT 26
AC134236/c 149349 bp DNA linear INV 21-DEC-2002
LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 3 clone
AC134236
AC134236.3 GI:27356674
KEYWORDS
HTG.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhacridae; Oryzaceae; Oryza.
1 (bases 1 to 149349)
Wing, R.A., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Currie, J. and
Collura, K.
Rice Genomic Sequence
Unpublished
2 (bases 1 to 149349)
Wing, R.A., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Sasaki, C.,
Currie, J. and Collura, K.
Direct Submission
Submitted (24-SEP-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
3 (bases 1 to 149349)
Wing, R.A., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Sasaki, C.,
Currie, J. and Collura, K.
Direct Submission
Submitted (23-OCT-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
4 (bases 1 to 149349)
Wing, R.A., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Currie, J. and
Collura, K.
Direct Submission
Submitted (21-DEC-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
On Dec 21, 2002 this sequence version replaced gi:24270659.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality
>=30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by more than
one plasmid subclone; and the assembly was confirmed by
restriction digest. There are three areas of single-subclones:
54466-54560, 52568-52653 and 115543-115568. From base 133735-
133753 there are only PCR reads. The assembly overlaps from base
1-22123 with OSJNBa0090010 (accession # AC134237). The overlap is
from 139009-162132 bases on OSJNBa0090010. The assembly overlaps
from base 136321-149349 with OJ1263H11 (accession # AC118980). The

FEATURES
source
1. 149349
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OSJNBa0056G13"
/note="(japonica cultivar-group)"

BASE COUNT
39620 a 33572 c 33754 g 42403 t
ORIGIN

Query Match 34.6%; Score 18; DB 8; Length 149349;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 TCGCTCTGCTGCTGCGG 40
|||||
Db 32462 TCGCTCTGCTGCTGCGG 32445

RESULT 27
AC009353/c 157371 bp DNA linear INV 14-DEC-2001
LOCUS
DEFINITION Drosophila melanogaster, chromosome 2R, region 41E-41E, BAC clone
BACR32F22, complete sequence.
AC009353
AC009353.8 GI:17737033
KEYWORDS
HTG.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 157371)
Celniker, S.E., Adams, M.D., Kronmiller, B., Tyler, D., Wan, K.H.,
Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C.,
Rogers, Y., An, H., Baldwin, D., Bansan, J., Beeson, K.Y., Busam, D.A.,
Carlson, J.W., Center, A., Champe, M., Davenport, L.B., Dietz, S.M.,
Dodson, K., Dorsett, V., Dou, L.E., Doyle, C., Drenek, D., Farfan, D.,
Fiertera, S., Frise, E., Galle, R.F., Garay, N.S., George, R.A.,
Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J.,
Ibegwam, C., Jalali, M., Kruse, D., Li, P., Matzel, B., Moshrefi, A.,
McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J.,
Paclet, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B.,
Phonathanavong, S., Plitman, G.S., Puri, V., Richards, S., Scheeler, F.,
Stapleton, M., Strong, R., Svrtkac, R., Tector, C., Williams, S.M.,
Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.
Sequencing of Drosophila chromosome 2R, region 41E-41E
Unpublished (1998)
2 (bases 1 to 157371)
Celniker, S.E., Ashbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Butenhoff, C., Champe, M., Chavez, C., Chew, M., Cleasithka, L.,
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
Hoskins, R.A., Houston, K.A., Hummachi, S.R., Katta, K., Kearney, L.,
Kim, E., Lee, B., Lewis, S., Li, P., Lomocan, M.A., Mazda, P.,
Moshrefi, A.R., Moshrefi, M., Nixon, K., Paclet, J.M., Park, S.,
Pfeiffer, B., Poon, L., Segueira, A., Sethi, H., Snir, E.,
Svrtkac, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and
Rubin, G.M.
Direct Submission
Submitted (17-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Dec 14, 2001 this sequence version replaced gi:1264782.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location

and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.

FEATURES

location/Qualifiers
1. 157371

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2R"
/map="41F-41E"
/clone="BACR32F22 (D910)"
/clone_1id="RPCL-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
PBAC3.6)"
BASE COUNT 47749 a 29939 c 31348 g 48335 t
ORIGIN

Query Match 34.6%; Score 18; DB 3; Length 157371;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 21 CATCGCTCTGCTGCTGC 38
Db 40255 CATGCTCTGCTGCTGC 40238

RESULT 28
AP003935
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 6 clone P0661G04,
*** SEQUENCING IN PROGRESS ***.

ACCESSION
AP003935
VERSION
KEYWORDS
SOURCE
ORGANISM
HTG: HTGS PHASE2.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
1

Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
clone: P0661G04
Published Only in Database (2001)

JOURNAL
AUTHORS
TITLE
JOURNAL
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (18-JUL-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kamondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: <http://rpg.dna.affrc.go.jp/>,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)

COMMENT

NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
location/Qualifiers

FEATURES

source

1. 159581
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="6"
/clone="P0661G04"

BASE COUNT 43622 a 35810 c 34889 g 44410 t 850 others
ORIGIN

Query Match 34.6%; Score 18; DB 2; Length 159581;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 17 CTCATCGCTCTGCTGC 34
Db 56020 CTCATCGCTCTGCTGC 56037

RESULT 29
AC119676
LOCUS
DEFINITION
AC119676 AL157905
ACCESSION
AC119676.2 GI:22024581
HTG.
KEYWORDS
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 160573)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
and Haugen, E.D.
Direct Submission

Unpublished
2 (bases 1 to 160573)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Direct Submission

Submitted (30-APR-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 160573)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
and Haugen, E.D.
Direct Submission

Submitted (31-JUL-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Jul 31, 2002 this sequence version replaced gi:20340501.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwgchgs@u.washington.edu
Dialing Center: SC

----- Project Information
Center Project name: chr-1
Center clone name: RP5-820016 (sc0838)

----- Summary Statistics
Sequencing vector: plasmid; 6% of reads
Sequencing vector: plasmid; L08752; 31% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 16046 bases at least Q40
Consensus quality: 16055 bases at least Q30
Consensus quality: 160573 bases at least Q20
Insert size: 160573; sum-of-contigs
Quality coverage: 8.7x in Q20 bases; sum-of-contigs

----- Overlapping Sequences:
5': RP11-486810 AL445933, 2000-bp overlap
3': RP11-216M14 AL451006, 2370-bp overlap

----- Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector, in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp)
are not resolved in the fingerprint and hence do not appear
in the table. There are no significant remaining discrepancies
between the experimental and predicted values. Uniquely ordered
fragments are separated by dashed lines.

HindIII

ECORI

BglIII

SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt

4651	4729	3216	3280	8871	8725
449	<800	2184	2124	5671	5578
512	<800	8065	7987	7274	7290
2814	2828	486	<800	5062	5218
1247	1223	106	<800	3738	3629
4695	4729	2252	2235	4303	4233
3373	3434	3261	3280	2976	2947
3480	3434	1176	1161	3575	3629
1040	1047	485	<800	4090	4059
4205	4315	730	760	4541	4505
1363	1327	1441	1415	5061	5015
2834	2828	8427	8263	1543	1500
580	<800	2059	2019	2059	2069
960	937	488	<800	1452	1438
345	<800	4826	4743	2922	2947
2226	2155	1007	996	793	<800
3655	3625	2022	2019	8159	7991
4826	4729	11962	11922	6687	6701
3185	3237	41	<800	3091	3122
1881	1846	121	<800	571	<800
1499	1482	318	<800	5239	5015
3571	3625	3966	3950	1981	1963
36	<800	6804	6606	272	<800
830	851	130	<800	2227	2226

1257	1223	6623	6606	1692	1722
769	781	7207	7242	11429	11239
1232	1223	6036	5928	15433	15722
3393	3434	2941	2941	5568	5578
2494	2506	9156	9188	8270	8296
4816	4729	2372	2367	11452	11239
367	<800	1292	1256	1221	1219
1707	1653	3253	3280	4183	4233
22	<800	770	760	453	<800
7538	7498	939	996	2058	2069
46	<800	1787	1737	2480	2497
14217	13799	1258	1256	9539	9559
6994	7061	6276	6300	1029	1028
2797	2828	291	<800	4901	5015
2801	2828	9501	9445	4717	4699
1858	1846	1001	996		
6740	6779	7642	7642		
476	<800	6018	5928		
11525	11253	6277	6300		
601	<800	6645	6606		
1236	1223	2105	2124		
4323	4315	3271	3280		
1630	1653	823	828		
1786	1846	528	<800		
2140	2155	1730	1737		
491	<800	7201	7242		
19364	19745	3020	3076		
14536	14543	3533	3528		
2699	2828	1514	1472		
2120	2155				
351	<800				

FEATURES
source Location/Qualifiers
1.160573
/organism="Homo sapiens"

Query Match 34.6% Score 18; DB 9; Length 160573;
Best Local Similarity 100.0%; Pred.No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GCAGTCTCTCAGCCTCT 20
|||||

DB 143869 GCAGTTCCTCAGCTCT 143886

RESULT 30
AC012389
LOCUS
DEFINITION
AC012389 161518 bp DNA linear HTG 24-FEB-2000
Drosophila melanogaster chromosome 3 clone BACR28N05 (D1183)
RFLP: 98.28 N.5 map 100F1-100F5 strain Y; cn bw sp, *** SEQUENCING
IN PROGRESS *** 70 unordered pieces.

ACCESSION
AC012389
VERSION
AC012389.10 GI:7025685
KEYWORDS
HTG: HTG_Phasel.
SOURCE
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
1 (bases 1 to 161518)
Celisner, S.E., Agdayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Butenhofer, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
Hinkle, A., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K.,
Kearney, L., Lee, B., Lewis, S., Li, P., Ling, H., Moshrefi, A.R.,
Moshrefi, M., Nixon, K., Paclebd, J.M., Park, S., Pfeiffer, B.,
Richard, S., Sethi, H., Svitskas, R.R., Wan, K.H., Webster, D.,
Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G.M.
Sequencing of Drosophila melanogaster
Unpublished

TITLE
JOURNAL
REFERENCE
AUTHORS
2 (bases 1 to 161518)
Celisner, S.E., Agdayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Butenhofer, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L.,
Kim, E., Lee, B., Lewis, S., Li, P., Lomocan, M.A., Mazda, P.,
Moshrefi, A.R., Moshrefi, M., Nixon, K., Paclebd, J.M., Park, S.,
Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Sutr, E.,
Svitskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and
Rubin, G.M.
Direct Submision
Submitted (27-OCT-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Feb 24, 2000 this sequence version replaced gi:6996991.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 70 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

TITLE
JOURNAL
COMMENT
1
569: contig of 569 bp in length
570
649: gap of unknown length
650
928: contig of 279 bp in length
929
1008: gap of unknown length
1009
1403: contig of 395 bp in length
1404
1483: gap of unknown length
1484
2103: contig of 620 bp in length
2104
2183: gap of unknown length
2184
2781: contig of 598 bp in length
2782
2861: gap of unknown length
2862
3522: contig of 661 bp in length
3523
3602: gap of unknown length
3603
4192: contig of 590 bp in length
4193
4272: gap of unknown length
4273
4898: contig of 626 bp in length
4899
4978: gap of unknown length
4979
5329: contig of 351 bp in length
5330
5409: gap of unknown length

5410
5954
5954: contig of 544 bp in length
5954
6033: gap of unknown length
6034
6948: contig of 915 bp in length
6949
7028: gap of unknown length
7029
8873: contig of 1845 bp in length
8874
8953: gap of unknown length
8954
10161: contig of 1208 bp in length
10162
10241: gap of unknown length
10242
12996: contig of 2755 bp in length
12997
13076: gap of unknown length
13077
14725: contig of 1649 bp in length
14726
14805: gap of unknown length
14806
16262: contig of 1457 bp in length
16263
16342: gap of unknown length
16343
18474: contig of 2132 bp in length
18475
18554: gap of unknown length
18555
20724: contig of 2170 bp in length
20725
20804: gap of unknown length
20805
24209: contig of 3404 bp in length
24209
24288: gap of unknown length
24289
26484: contig of 2116 bp in length
26485
26485: gap of unknown length
26485
29108: contig of 2624 bp in length
29108
32550: gap of unknown length
32550
32630: contig of 3362 bp in length
32631
35639: gap of unknown length
35639
35719: contig of 3006 bp in length
35720
35720: gap of unknown length
35720
38937: contig of 3218 bp in length
38938
39017: gap of unknown length
39017
42731: contig of 3713 bp in length
42731
42810: gap of unknown length
42811
47071: contig of 4261 bp in length
47072
47151: gap of unknown length
47152
51484: contig of 4333 bp in length
51485
51564: gap of unknown length
51565
55312: contig of 3748 bp in length
55313
55392: gap of unknown length
55393
59360: contig of 3968 bp in length
59361
59440: gap of unknown length
59441
67279: contig of 7839 bp in length
67280
67359: gap of unknown length
67360
71454: contig of 4095 bp in length
71455
71534: gap of unknown length
71535
77682: contig of 6068 bp in length
77683
77683: gap of unknown length
84383: contig of 6701 bp in length
84384
84463: gap of unknown length
84464
92633: contig of 8170 bp in length
92634
92713: gap of unknown length
92714
100750: contig of 8037 bp in length
100751
100830: gap of unknown length
100831
112677: contig of 11647 bp in length
112678
112757: gap of unknown length
112758
121914: contig of 9157 bp in length
121915
121994: gap of unknown length
121995
139643: contig of 17649 bp in length
139644
139723: gap of unknown length
139723
140359: contig of 636 bp in length
140360
140439: gap of unknown length
140440
141138: contig of 699 bp in length
141139
141218: gap of unknown length
141219
141882: contig of 764 bp in length
141883
142062: gap of unknown length
142063
143327: contig of 1265 bp in length
143328
143407: gap of unknown length
143408
144194: contig of 787 bp in length
144195
144374: gap of unknown length
144375
145920: contig of 746 bp in length
145921
145100: gap of unknown length
145101
145588: contig of 488 bp in length
145589
145668: gap of unknown length
145669
146311: contig of 643 bp in length

146312 146391: gap of unknown length
 146382 146960: contig of 569 bp in length
 146961 147040: gap of unknown length
 147041 147664: contig of 624 bp in length
 147665 147744: gap of unknown length
 147745 148437: contig of 693 bp in length
 148438 148517: gap of unknown length
 148518 148960: contig of 443 bp in length
 148961 149040: gap of unknown length
 149041 149606: contig of 566 bp in length
 149607 149687: gap of unknown length
 149688 150570: contig of 884 bp in length
 150571 150650: gap of unknown length
 150651 151251: contig of 601 bp in length
 151252 151331: gap of unknown length
 151332 152032: contig of 701 bp in length
 152033 152112: gap of unknown length
 152113 152660: contig of 548 bp in length
 152661 152740: gap of unknown length
 152741 152983: contig of 243 bp in length
 152984 153063: gap of unknown length
 153064 153312: contig of 249 bp in length
 153313 153392: gap of unknown length
 153393 154241: contig of 849 bp in length
 154242 154321: gap of unknown length
 154322 154712: contig of 391 bp in length
 154713 154792: gap of unknown length
 154793 155639: contig of 847 bp in length
 155640 155719: gap of unknown length
 155720 156286: contig of 567 bp in length
 156287 156366: gap of unknown length
 156367 156617: contig of 248 bp in length
 156618 156694: gap of unknown length
 156695 157047: contig of 353 bp in length
 157048 157127: gap of unknown length
 157128 157560: contig of 433 bp in length
 157561 157640: gap of unknown length
 157641 158264: contig of 624 bp in length
 158265 158344: gap of unknown length
 158345 158951: contig of 607 bp in length
 158952 159031: gap of unknown length
 159032 159561: contig of 530 bp in length
 159562 159641: gap of unknown length
 159642 160168: contig of 527 bp in length
 160169 160248: gap of unknown length
 160249 160801: contig of 553 bp in length
 160802 160881: gap of unknown length
 160882 161518: contig of 637 bp in length.

FEATURES
 source
 1. 161518
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /strain="y; cn bw sp"
 /db_xref="taxon:7227"
 /chromosome="3"
 /map="100F1-100F5"
 /clone="BACR28N05 (D1183) RPCT-98 28.N.5"
 /clone_1lb="RPCT-98 (Roswell Park Cancer Institute
 Drosophila melanogaster BAC library, partial EcoRI in

Query Match 34.6%; Score 18; DB 2; Length 161518;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 CATCGCTCTGCTGCTGC 38
 Db 86402 CATCGCTCTGCTGCTGC 86419

RESULT 31
 AC013579/c 173093 bp DNA linear HTG 12-MAR-2000
 LOCUS Homo sapiens clone RP11-3C4, WORKING DRAFT SEQUENCE, 26 unordered

ACCESSION
 AC013579
 VERSION
 AC013579.4 GI:7229999
 HTG; HTGS PHASE1; HTGS_DRAFT.
 KEYWORDS
 Homo sapiens (human)
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 1 (bases 1 to 173093)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens, clone RP11-3C4
 Unpublished
 2 (bases 1 to 173093)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
 Baldin,J., Barina,N., Beckert,R., Boguslavsky,L., Boukhalter,B.,
 Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
 Cooke,P., DeArillano,K., Dewar,K., Domingo,M., Donelan,L., Doyle,M.,
 Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J.,
 Galagan,J.C., Gardyna,S., Grant,G., Hagos,B., Heathord,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 Lehoczy,J., Liu,C., Locke,K., Macdonald,P., Marquis,N.,
 McEwan,P., McGurk,A., McKernan,K., McLoughlin,J., Meidrim,D.,
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
 Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Testaye,S., Tittrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
 Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 12, 2000 this sequence version replaced gi:6984402.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE
 JOURNAL
 COMMENT
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L2712
 Center clone name: 3 C 4

Summary Statistics
 Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 158945 bases at least Q40
 Consensus quality: 164943 bases at least Q30
 Consensus quality: 168064 bases at least Q20
 Insert size: 182000; agarose-fp
 Insert size: 170593; sum-of-contigs
 Quality coverage: 3.7 in Q20 bases; agarose-fp
 Quality coverage: 3.9 in Q20 bases; sum-of-contigs
 NOTE: This is a 'working draft' sequence. It currently
 consists of 26 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1 1347: contig of 1347 bp in length
 1348 1447: gap of 100 bp
 1448 2769: contig of 1322 bp in length
 2770 2870 2870: gap of 100 bp
 2870 4168: contig of 1299 bp in length
 4169 4268: gap of 100 bp
 4269 6550: contig of 2282 bp in length
 6551 6650: gap of 100 bp
 6651 8819: contig of 2169 bp in length
 8820 8919: gap of 100 bp
 8920 10822: contig of 1903 bp in length


```
* 10823 10922: gap of 100 bp
* 10923 13176: contig of 2254 bp in length
* 13177 13276: gap of 100 bp
* 13277 15574: contig of 2298 bp in length
* 15575 15674: gap of 100 bp
* 15675 17611: contig of 1937 bp in length
* 17612 17712: gap of 100 bp
* 17712 20922: contig of 3211 bp in length
* 20923 21022: gap of 100 bp
* 21023 24551: contig of 3529 bp in length
* 24552 24652: gap of 100 bp
* 24652 28391: contig of 3739 bp in length
* 28391 28491: gap of 100 bp
* 28491 33813: contig of 5323 bp in length
* 33814 33913: gap of 100 bp
* 33914 37959: contig of 4046 bp in length
* 37960 38059: gap of 100 bp
* 38060 42630: contig of 4571 bp in length
* 42631 42730: gap of 100 bp
* 42731 50918: contig of 8188 bp in length
* 50919 51018: gap of 100 bp
* 51019 59273: contig of 8255 bp in length
* 59274 59373: gap of 100 bp
* 59374 68691: contig of 9318 bp in length
* 68692 68791: gap of 100 bp
* 68792 76416: contig of 7625 bp in length
* 76417 76516: gap of 100 bp
* 76517 83772: contig of 7256 bp in length
* 83773 83872: gap of 100 bp
* 83873 93356: contig of 9484 bp in length
* 93357 93456: gap of 100 bp
* 93457 105377: contig of 11921 bp in length
* 105378 105477: gap of 100 bp
* 105478 118368: contig of 12891 bp in length
* 118369 118468: gap of 100 bp
* 118469 135438: contig of 16970 bp in length
* 135439 135538: gap of 100 bp
* 135539 154849: contig of 19311 bp in length
* 154850 154949: gap of 100 bp
* 154950 173093: contig of 18144 bp in length.
```

FEATURES

```
source
1. 173093
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /clone="RP11-3C4"
  /clone_lib="RP11-11 Human Male BAC"
  1. 1347
    /note="assembly_fragment"
    1448. 2769
      /note="assembly_fragment"
      2870. 4168
        /note="assembly_fragment"
        4269. 6550
          /note="assembly_fragment"
          6651. 8819
            /note="assembly_fragment"
            8920. 10822
              /note="assembly_fragment"
              10923. 13176
                /note="assembly_fragment"
                13277. 15574
                  /note="assembly_fragment"
                  15675. 17611
                    /note="assembly_fragment"
                    17712. 20922
                      /note="assembly_fragment"
                      21023. 24551
                        /note="assembly_fragment"
                        24652. 28390
                          /note="assembly_fragment"
                          28491. 33813
                            /note="assembly_fragment"
```

```
misc_feature 33914..37959
              /note="assembly_fragment"
              38060..42630
                /note="assembly_fragment"
                42731..50918
                  /note="assembly_fragment"
                  51019..59273
                    /note="assembly_fragment"
                    59374..68691
                      /note="assembly_fragment"
                      68792..76416
                        /note="assembly_fragment"
                        clone_end:"7"
                        vector_side:"left"
                        76517..83772
                          /note="assembly_fragment"
                          clone_end:"SP6"
                          vector_side:"right"
                          83873..93356
                            /note="assembly_fragment"
                            93457..105377
                              /note="assembly_fragment"
                              105478..118368
                                /note="assembly_fragment"
                                118469..135438
                                  /note="assembly_fragment"
                                  135539..154849
                                    /note="assembly_fragment"
                                    154950..173093
                                      /note="assembly_fragment"
```

```
BASE COUNT 45233 a 40721 c 40396 g 44224 t 2519 others
ORIGIN
```

```
Query Match 34.6%; Score 18; DB 2; Length 173093;
Best Local Similarity 100.0%; Pred. No. 36;
```

```
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Cy 3 GCAGTCTCTACCGCTCT 20
```

```
Db 168719 GCAGTCTCTACCGCTCT 168702
```

```
RESULT 32
AE003845
LOCUS 173874 bp DNA linear INV 14-FEB-2003
DEFINITION Drosophila melanogaster chromosome 4 section 1 of 5 of the complete
sequence.
ACCESSION AE003845 AE014135
VERSION AE003845.4 GI:28380223
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 173874)
Adams,M.D., Celisner,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
Amaratides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galie,R.F.,
George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
Brandon,R.C., Rogers,Y.H., Blazer,R.G., Champe,M., Pfeiffer,B.D.,
Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabot,G.L.,
Abail,J.F., Agbayani,A., An,H.J., Andrews-Plamkoch,C., Baldwin,D.,
Ballew,R.M., Baer,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M.,
Beeson,K.Y., Benos,P.V., Bernan,B.P., Bhandari,D., Bolshakov,S.,
Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brothier,P.,
Burtis,K.C., Buzam,D.A., Butler,H., Cadieu,E., Center,A.,
Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B.,
Davies,P., de Pablos,B., Delcher,A., Deng,Z., Dwyer,A.D., Dew,I.,
Dietz,S.M., Dodson,K., Doup,J.E., Downes,M., Dugan-Rocha,S.,
Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C.,
Ferreira,S., Fleischmann,W., Fowler,C., Gabrielian,A.E., Garg,N.S.,
Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
```



```

CDS
    /locus_tag="CG17923"
    /product="CG17923-PA"
    /db_xref="FLYBASE:FBgn0040037"
    join(24134..24477,24979..25153,25218..25450,25501..25621)
    /locus_tag="CG17923"
    /note="CG17923 gene product"
    /codon_start=1
    /product="CG17923-PA"
    /protein_id="AAN06537.1"
    /db_xref="GI:22759416"
    /db_xref="FLYBASE:FBgn0040037"
    /translation="MGISSDYSKQADADILDDNFASIVGIEGRILFDNLKKSIA
    YLTLSNLPETVPLFLFVFDIPALGTAALICIDIGTDLPLPISLAYEAKEDIDAM
    PRDPEDRLVNNKLLFMVLIQIVGTACFTFPAIMHGHFPSPRLKIGEDMSK
    NVEDLEDGGEEMTVRERKRVLETAGTGFEVSIVTVQVFDLICRNRSLIOGCMG
    HVLPALVLEFIACLLCVPFVEKTLNMYSIKFIWIAVAFGLLIFFPDSRFLI
    RNRPGWVBOERY"
    complement(26482..34110)
    /locus_tag="CG32011"
    /note="1st curated on Thu Aug 29 10:39:29 PDT 2002"
    /map="102A1-102A1"
    /db_xref="FLYBASE:FBgn0052011"
    complement(join(26482..26661,27167..27349,28371..28609,
    28966..29301,29356..30391,30551..31625,31703..33391,
    33949..34110))

Query Match      34.6%  Score 18; DB 3; Length 173874;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      21 CATCGCTCTGCTGCTGC 38
      |||||
Db      12415 CATCGCTCTGCTGCTGC 12432

RESULT 33
AC007587      185916 bp      DNA      linear      INV 14-DEC-2001
LOCUS      Drosophila melanogaster, chromosome 2R, region 41C-41D, BAC clone
AC007587
VERSION      AC007587.11      GI:17737031
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 185916)
Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brannon,R.C.,
Rogers,Y., An,H., Baldwin,D., Bantzou,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Chape,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dreier,D., Fartan,D.,
Fertler,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houch,J., Hoskins,R.A., Hoslin,D., Howland,T.J.,
Ibegan,C., Jalili,M., Kruse,D., Li,P., Matrei,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunco,J.,
Paclob,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
Phuanavanong,S., Pitterman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svitskas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M., and Venter,J.C.
Sequencing of Drosophila chromosome 2R, region 41C-41D
Unpublished (1998)

TITLE
JOURNAL
REFERENCE
AUTHORS

```

```

TITLE
JOURNAL
COMMENT
Rubin,G.M.
Direct Submission
Submitted (19-MAY-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Dec 14, 2001 this sequence version replaced gi:7264774.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu.
Location/Qualifiers
1..185916
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2R"
/map="41C-41D"
/clone_1b="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcorI in
pBAC3.6)"

BASE COUNT      56016 a 37375 c 36026 g 56499 t
ORIGIN

Query Match      34.6%  Score 18; DB 3; Length 185916;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      21 CATCGCTCTGCTGCTGC 38
      |||||
Db      29892 CATCGCTCTGCTGCTGC 29909

RESULT 34
AC008333      187223 bp      DNA      linear      HTG 15-MAR-2002
LOCUS      Drosophila melanogaster chromosome 2 clone BACR02C02 (D869) RPCI-98
AC008333
DEFINITION      *** 117 unordered pieces.
AC008333
VERSION      AC008333.22      GI:19482330
KEYWORDS      HTG; HTGS PHASE1.
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 187223)
Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Buenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Fartan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Paclob,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Shit,E.,
Svitskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
Sequencing of Drosophila melanogaster
Unpublished

TITLE
JOURNAL
REFERENCE
AUTHORS

```

TITLE	JOURNAL	REFERENCE	AUTHORS
Rubin,G.M.	Direct Submission	Submitted (02-AUG-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA	3 (bases 1 to 187223)
Celniker,S.E., Agbayani,A., Arcaine,T.T., Baxter,E., Blazej,R.G., Butenoff,C., Champe,M., Chavez,C., Chew,M., Cieciolka,L., Doyle,C.M., Faltan,D.E., Galle,R., George,R.A., Harris,N.L., Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karta,K., Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacley,J.M., Park,S., Pfeiffer,B., Richards,S., Sethi,H., Swirskas,R.R., Tan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.	Direct Submission	Submitted (15-MAR-2002) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA	On Mar 15, 2002 this sequence version replaced g1:7264781. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (http://www.fruitfly.org/sequence/) or send email to bdgs@fruitfly.berkeley.edu . All contigs in this submission meet the following cutoffs: length >= 200 bases.
NOTE: This is a 'working draft' sequence. It currently consists of 117 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.			
1	399: contig of 399 bp in length		
400	479: gap of unknown length		
480	1585: contig of 1106 bp in length		
1586	1665: gap of unknown length		
1666	3031: contig of 1366 bp in length		
3032	3111: gap of unknown length		
3112	4483: contig of 1372 bp in length		
4484	4563: gap of unknown length		
4564	5967: contig of 1404 bp in length		
5968	6047: gap of unknown length		
6048	8158: contig of 2111 bp in length		
8159	8238: gap of unknown length		
8239	9368: contig of 1130 bp in length		
9369	9448: gap of unknown length		
9449	11398: contig of 1950 bp in length		
11399	11478: gap of unknown length		
11479	12823: contig of 1345 bp in length		
12824	12903: gap of unknown length		
12904	15331: contig of 2428 bp in length		
15332	15411: gap of unknown length		
15412	16843: contig of 1432 bp in length		
16844	16923: gap of unknown length		
16924	19239: contig of 2316 bp in length		
19240	19319: gap of unknown length		
19320	20191: contig of 872 bp in length		
20192	20271: gap of unknown length		
20272	23177: contig of 2906 bp in length		
23178	23257: gap of unknown length		
23258	24861: contig of 1604 bp in length		
24862	24941: gap of unknown length		
24942	28264: contig of 3323 bp in length		
28265	28344: gap of unknown length		
28345	32799: contig of 4455 bp in length		
32800	32879: gap of unknown length		
32880	44930: contig of 12051 bp in length		
44931	45010: gap of unknown length		
45011	59268: contig of 14258 bp in length		
59269	59348: gap of unknown length		
59349	73716: contig of 14368 bp in length		
73717	73796: gap of unknown length		
73797	85947: contig of 12151 bp in length		
85948	86027: gap of unknown length		
86028	107333: contig of 21306 bp in length		

107334	107413: gap of unknown length		
107414	116139: contig of 8726 bp in length		
116140	116219: gap of unknown length		
116220	130488: contig of 14269 bp in length		
130489	130568: gap of unknown length		
130569	130881: contig of 313 bp in length		
130882	130961: gap of unknown length		
130962	131434: contig of 473 bp in length		
131435	131514: gap of unknown length		
131515	132452: contig of 938 bp in length		
132453	132532: gap of unknown length		
132533	133533: contig of 1001 bp in length		
133534	133613: gap of unknown length		
133614	134016: contig of 403 bp in length		
134017	134096: gap of unknown length		
134097	134556: contig of 460 bp in length		
134557	134636: gap of unknown length		
134637	135518: contig of 882 bp in length		
135519	135598: gap of unknown length		
135599	136290: contig of 692 bp in length		
136291	136370: gap of unknown length		
136371	136924: contig of 554 bp in length		
136925	137004: gap of unknown length		
137005	137715: contig of 711 bp in length		
137716	137795: gap of unknown length		
137796	138094: contig of 299 bp in length		
138095	138174: gap of unknown length		
138175	139004: contig of 830 bp in length		
139005	139084: gap of unknown length		
139085	139649: contig of 565 bp in length		
139650	139729: gap of unknown length		
139730	140353: contig of 624 bp in length		
140354	140433: gap of unknown length		
140434	140639: contig of 206 bp in length		
140640	140719: gap of unknown length		
140720	141001: contig of 282 bp in length		
141002	141081: gap of unknown length		
141082	142004: contig of 923 bp in length		
142005	142084: gap of unknown length		
142085	142643: contig of 559 bp in length		
142644	142723: gap of unknown length		
142724	143061: contig of 338 bp in length		
143062	143141: gap of unknown length		
143142	143398: contig of 257 bp in length		
143399	143478: gap of unknown length		
143479	144700: contig of 1222 bp in length		
144701	144780: gap of unknown length		
144781	145010: contig of 230 bp in length		
145011	145090: gap of unknown length		
145091	145783: contig of 693 bp in length		
145784	145863: gap of unknown length		
145864	146308: contig of 445 bp in length		
146309	146388: gap of unknown length		
146389	147049: contig of 661 bp in length		
147050	147129: gap of unknown length		
147130	147551: contig of 422 bp in length		
147552	147631: gap of unknown length		
147632	148164: contig of 533 bp in length		
148165	148244: gap of unknown length		
148245	148503: contig of 259 bp in length		
148504	148583: gap of unknown length		
148584	149079: contig of 496 bp in length		
149080	149159: gap of unknown length		
149160	150009: contig of 850 bp in length		
150010	150089: gap of unknown length		
150090	151007: contig of 918 bp in length		
151008	151087: gap of unknown length		
151088	151607: contig of 520 bp in length		
151608	151687: gap of unknown length		
151688	152125: contig of 438 bp in length		
152126	152205: gap of unknown length		
152206	152550: contig of 345 bp in length		
152551	152630: gap of unknown length		

```

* 152631 153442: contig of 712 bp in length
* 153443 153422: gap of unknown length
* 153423 154116: contig of 694 bp in length
* 154117 154196: gap of unknown length
* 154197 154542: contig of 446 bp in length
* 154643 154722: gap of unknown length
* 154723 155046: contig of 324 bp in length
* 155047 155126: gap of unknown length
* 155127 156443: contig of 1317 bp in length
* 156444 156523: gap of unknown length
* 156524 157098: contig of 576 bp in length
* 157099 157178: gap of unknown length
* 157179 157983: contig of 805 bp in length
* 157984 158063: gap of unknown length
* 158064 159681: contig of 1618 bp in length
* 159682 159761: gap of unknown length
* 159762 160166: contig of 405 bp in length
* 160167 160246: gap of unknown length
* 160247 160579: contig of 331 bp in length
* 160580 160659: gap of unknown length
* 160660 160963: contig of 304 bp in length
* 160964 161043: gap of unknown length

```

```

Query Match      34.6%; Score 18; DB 2; Length 187223;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      21 CATGCGTCTCGTGGCTGC 38
Db      129760 CATGCGTCTCGTGGCTGC 129743

```

```

RESULT 35
AC016393/c
LOCUS      Homo sapiens chromosome 12 clone RP11-199C7 map 12, *** SEQUENCING
DEFINITION
IN PROGRESS **, 91 unordered pieces.
ACCESSION  AC016393
VERSION     AC016393.3
KEYWORDS    HTG: HTGS, PHASE1.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 188032)
AUTHORS   Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE      Homo sapiens chromosome 12, clone RP11-199C7
JOURNAL    Unpublished
REFERENCE 2 (bases 1 to 188032)
AUTHORS   Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
          Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhalter, B.,
          Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
          Cooke, P., Dearrellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
          Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
          Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
          Howland, J., Lieu, C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
          Lehoczeky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
          McMan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
          Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
          Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
          Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
          Testave, S., Titrrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
          Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

```

```

TITLE      Direct Submission
JOURNAL    Submitted (25-NOV-1999) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
          3 (bases 1 to 188032)
REFERENCE 3 (bases 1 to 188032)
AUTHORS   Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
          Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
          Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,
          Campoliano, A., Castle, A., Choepe, Y., Colangelo, M., Collins, S.,
          Collymore, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J., S.,
          Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,

```

TITLE
JOURNAL
COMMENT

Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J., C. Lieu, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Lacoque, K., Lamazares, R., Lander, E., Lehoczeky, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McMan, P., McGurk, A., McKernan, K., McPheters, R.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Olivier, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testave, S., Theodore, J., Titrrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, C., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 12, 2000 this sequence version replaced gi:6759157.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

Center project name: L4953

Center clone name: 199_C_7

* NOTE: This is a 'working draft' sequence. It currently
* consists of 91 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

1
1264 1263: contig of 1263 bp in length
1364 1363: gap of 100 bp
2365 2365: contig of 1002 bp in length
2465 2465: gap of 100 bp
3982 3982: contig of 1517 bp in length
4082 4082: gap of 100 bp
5089 5089: contig of 1007 bp in length
5189 5189: gap of 100 bp
5090 5090: gap of 100 bp
6231 6231: contig of 1042 bp in length
6332 6332: gap of 100 bp
6332 6332: gap of 100 bp
7581 7580: contig of 1249 bp in length
7681 7680: gap of 100 bp
8972 8972: contig of 1292 bp in length
9072 9072: gap of 100 bp
10251 10251: contig of 1179 bp in length
10351 10351: gap of 100 bp
11684 11684: contig of 1333 bp in length
11784 11784: gap of 100 bp
12954 12954: contig of 1170 bp in length
13054 13054: gap of 100 bp
14166 14166: contig of 1112 bp in length
14265 14265: gap of 100 bp
15381 15381: contig of 1115 bp in length
15382 15381: gap of 100 bp
15481 15481: gap of 100 bp
16485 16485: contig of 1004 bp in length
16486 16485: gap of 100 bp
17950 17950: contig of 1365 bp in length
17951 17951: gap of 100 bp
18050 18050: gap of 100 bp
18051 18050: contig of 12 bp in length
18051 18050: gap of 100 bp
18162 18162: gap of 100 bp
18163 18162: gap of 100 bp
19591 19591: contig of 1429 bp in length
19592 19591: gap of 100 bp
20869 20869: contig of 1178 bp in length
20870 20869: gap of 100 bp

```

```

20970 22151: contig of 1182 bp in length
22152 22251: gap of 100 bp
22252 23495: contig of 1244 bp in length
23496 23595: gap of 100 bp
23596 24843: contig of 1248 bp in length
24844 24943: gap of 100 bp
24944 26155: contig of 1212 bp in length
26156 26255: gap of 100 bp
26256 27991: contig of 1736 bp in length
27992 28091: gap of 100 bp
28092 29183: contig of 1092 bp in length
29184 29283: gap of 100 bp
29284 30283: contig of 1010 bp in length
30294 30393: gap of 100 bp
30394 31466: contig of 1073 bp in length
31467 31566: gap of 100 bp
31567 32928: contig of 1362 bp in length
32929 33028: gap of 100 bp
33029 34776: contig of 1748 bp in length
34777 34876: gap of 100 bp
34877 36265: contig of 1389 bp in length
36266 36365: gap of 100 bp
36366 37605: contig of 1240 bp in length
37606 37705: gap of 100 bp
37706 38974: contig of 1269 bp in length
38975 39074: gap of 100 bp
39075 40313: contig of 1239 bp in length
40314 40413: gap of 100 bp
40414 41544: contig of 1131 bp in length
41545 41644: gap of 100 bp
41645 43458: contig of 1814 bp in length
43459 43558: gap of 100 bp
43559 45089: contig of 1531 bp in length
45090 45189: gap of 100 bp
45190 46414: contig of 1225 bp in length
46415 46514: gap of 100 bp
46515 47776: contig of 1262 bp in length
47777 47876: gap of 100 bp
47877 49306: contig of 1430 bp in length
49307 49406: gap of 100 bp
49407 50583: contig of 1177 bp in length
50584 50683: gap of 100 bp
50684 52614: contig of 1931 bp in length
52615 52714: gap of 100 bp
52715 54063: contig of 1349 bp in length
54064 54163: gap of 100 bp
54164 55229: contig of 1066 bp in length
55230 55329: gap of 100 bp
55330 56490: contig of 1161 bp in length
56491 56590: gap of 100 bp
56591 57831: contig of 1241 bp in length
57832 57931: gap of 100 bp
57932 59583: contig of 1652 bp in length
59584 59683: gap of 100 bp
59684 61686: contig of 2003 bp in length
61687 61786: gap of 100 bp
61787 63126: contig of 1340 bp in length
63127 63226: gap of 100 bp
63227 64917: contig of 1691 bp in length
64918 65017: gap of 100 bp
65019 66634: contig of 1617 bp in length
66635 66734: gap of 100 bp
66735 68500: contig of 1766 bp in length
68501 68600: gap of 100 bp
68601 70266: contig of 1666 bp in length
70267 70366: gap of 100 bp
70367 72119: contig of 1753 bp in length
72120 72219: gap of 100 bp
72220 74200: contig of 1981 bp in length
74201 74300: gap of 100 bp
74301 76382: contig of 2082 bp in length
76383 76482: gap of 100 bp
76483 78034: contig of 1552 bp in length

```

```

Query Match      34.6%  Score 18; DB 2; Length 188032;
Best Local Similarity 100.0%; Pred. NO. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 5 AGTTCTCTCAGCCTCTCA 22
Db 116725 AGTTCTCTCAGCCTCTCA 116708

```

```

RESULT 36
AC013254
LOCUS
DEFINITION
AC013254 189893 bp DNA linear HTG 17-MAR-2000
PCT-98.25.0.10 map 40D-40D strain Y; cn bw sp. *** SEQUENCING IN
PROGRESS ***; 184 unordered pieces.

```

```

ACCESSION
AC013254
VERSION
AC013254.6 GI:7259653
KEYWORDS
HTG; HTGS PHASE1.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

```

```

REFERENCE
1 (bases 1 to 189893)
Celiker,S.E., Aghayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champagne,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karta,K.,
Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacleb,D.M., Park,S., Pfeiffer,B.,
Richards,S., Sethi,H., Swirek,R.R., Wan,K.H., Webster,D.,
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
Sequencing of Drosophila melanogaster

```

```

TITLE
JOURNAL
REFERENCE
2 (bases 1 to 189893)
Celiker,S.E., Aghayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champagne,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karta,K.,
Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacleb,D.M., Park,S., Pfeiffer,B.,
Richards,S., Sethi,H., Swirek,R.R., Wan,K.H., Webster,D.,
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
Direct Submision

```

```

TITLE
JOURNAL
COMMENT
Submitted (05-NOV-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Mar 17, 2000 this sequence version replaced gi:5984327.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 184 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

```

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 588: contig of 588 bp in length
* 589 668: gap of unknown length
* 669 1063: contig of 395 bp in length
* 1064 1143: gap of unknown length
* 1144 1659: contig of 516 bp in length
* 1660 1739: gap of unknown length
* 1740 2258: contig of 519 bp in length
* 2259 2338: gap of unknown length
* 2339 2533: contig of 295 bp in length
* 2534 2713: gap of unknown length
* 2714 3238: contig of 525 bp in length
* 3239 3318: gap of unknown length
* 3319 3832: contig of 514 bp in length
* 3833 3912: gap of unknown length
* 3913 4239: contig of 327 bp in length
* 4240 4320: gap of unknown length
* 4320 4912: contig of 593 bp in length
* 4913 4992: gap of unknown length
* 4993 5610: contig of 618 bp in length
* 5611 5690: gap of unknown length
* 5691 5976: contig of 286 bp in length
* 5977 6056: gap of unknown length
* 6057 6580: contig of 524 bp in length
* 6581 7196: contig of 536 bp in length
* 7197 7276: gap of unknown length
* 7277 7861: contig of 585 bp in length
* 7862 7941: gap of unknown length
* 7942 9001: contig of 1060 bp in length
* 9002 9684: contig of 603 bp in length
* 9685 9764: gap of unknown length
* 9765 10668: contig of 904 bp in length
* 10669 10748: gap of unknown length
* 10749 11441: contig of 693 bp in length
* 11442 11521: gap of unknown length
* 11522 12133: contig of 612 bp in length
* 12134 12213: gap of unknown length
* 12214 13064: contig of 851 bp in length
* 13065 13144: gap of unknown length
* 13145 13935: contig of 791 bp in length
* 13936 14015: gap of unknown length
* 14016 14921: contig of 906 bp in length
* 14922 15001: gap of unknown length
* 15002 16065: contig of 1064 bp in length
* 16066 16145: gap of unknown length
* 16146 16953: contig of 808 bp in length
* 16954 17033: gap of unknown length
* 17034 17569: contig of 536 bp in length
* 17570 17649: gap of unknown length
* 17650 18097: contig of 448 bp in length
* 18098 18177: gap of unknown length
* 18178 18940: contig of 763 bp in length
* 18941 19020: gap of unknown length
* 19021 19247: contig of 227 bp in length
* 19248 19327: gap of unknown length
* 19328 20043: contig of 716 bp in length
* 20044 20124: gap of unknown length
* 20124 20734: contig of 611 bp in length
* 20735 20814: gap of unknown length
* 20815 21151: contig of 337 bp in length
* 21152 21231: gap of unknown length
* 21232 21885: contig of 654 bp in length
* 21886 21965: gap of unknown length
* 21966 23030: contig of 1065 bp in length
* 23031 23110: gap of unknown length
* 23111 24601: contig of 1491 bp in length
* 24602 24681: gap of unknown length
* 24682 25250: contig of 569 bp in length
* 25251 25330: gap of unknown length

* 25331 26463: contig of 1133 bp in length
* 26464 26543: gap of unknown length
* 26544 26937: contig of 394 bp in length
* 26938 27017: gap of unknown length
* 27018 28151: contig of 1134 bp in length
* 28152 28231: gap of unknown length
* 28232 29547: contig of 1316 bp in length
* 29548 29627: gap of unknown length
* 29628 30471: contig of 844 bp in length
* 30472 30551: gap of unknown length
* 30552 32369: contig of 1818 bp in length
* 32370 32449: gap of unknown length
* 32450 33789: contig of 1340 bp in length
* 33790 33869: gap of unknown length
* 33870 35224: contig of 1355 bp in length
* 35225 35304: gap of unknown length
* 35305 36553: contig of 1249 bp in length
* 36554 36633: gap of unknown length
* 36634 38274: contig of 1641 bp in length
* 38275 38354: gap of unknown length
* 38355 38389: contig of 1535 bp in length
* 38390 39969: gap of unknown length
* 39970 41731: contig of 1762 bp in length
* 41732 43829: contig of 2018 bp in length
* 43830 43909: gap of unknown length
* 43910 46534: contig of 2725 bp in length
* 46535 46714: gap of unknown length
* 46715 48565: contig of 1851 bp in length
* 48566 48645: gap of unknown length
* 48646 50553: contig of 1908 bp in length
* 50554 50633: gap of unknown length
* 50634 53539: contig of 3006 bp in length
* 53540 53719: gap of unknown length
* 53720 55306: contig of 1587 bp in length
* 55307 61387: contig of 6001 bp in length
* 61388 61467: gap of unknown length
* 61468 66422: contig of 4955 bp in length
* 66423 66502: gap of unknown length
* 66503 72908: contig of 6406 bp in length
* 72909 72988: gap of unknown length
* 72989 79928: contig of 6940 bp in length
* 79929 80008: gap of unknown length
* 80009 88702: contig of 8664 bp in length
* 88703 88782: gap of unknown length
* 88783 99255: contig of 10473 bp in length
* 99256 99335: gap of unknown length
* 99336 99914: contig of 579 bp in length
* 99915 99994: gap of unknown length
* 99995 100377: contig of 383 bp in length
* 100378 100457: gap of unknown length
* 100458 100944: contig of 487 bp in length
* 100945 101024: gap of unknown length
* 101025 101605: contig of 581 bp in length
* 101606 101685: gap of unknown length
* 101686 102540: contig of 855 bp in length
* 102541 102620: gap of unknown length
* 102621 103304: contig of 684 bp in length
* 103305 103384: gap of unknown length
* 103385 103818: contig of 434 bp in length
* 103819 103898: gap of unknown length
* 103899 104685: contig of 787 bp in length
* 104686 104765: gap of unknown length
* 104766 105637: contig of 872 bp in length
* 105638 105717: gap of unknown length
* 105718 106017: contig of 300 bp in length
* 106018 106097: gap of unknown length
* 106098 106348: contig of 251 bp in length
* 106349 106428: gap of unknown length
* 106429 107046: contig of 618 bp in length
* 107047 107126: gap of unknown length
* 107127 107624: contig of 498 bp in length


```
* 107625 107704: gap of unknown length
* 107705 108652: contig of 958 bp in length
* 108663 108742: gap of unknown length
* 108743 109478: contig of 736 bp in length
* 109479 109558: gap of unknown length
* 109559 110172: contig of 614 bp in length
* 110173 110252: gap of unknown length
* 110253 110986: contig of 734 bp in length

Query Match      34.6%; Score 18; DB 2; Length 189893;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 21 CATCGCTCTGCTGCTGC 38
Db 64905 CATCGCTCTGCTGCTGC 64922

RESULT 37
AC013254/c
LOCUS
DEFINITION
AC013254 189893 bp DNA linear HTG 17-MAR-2000
Drosophila melanogaster chromosome 2 clone BACR25010 (D1173)
PFC1-98.25.O.10 map 40D-40D strain Y; cn bw sp, *** SEQUENCING IN
PROGRESS ***. 184 unordered pieces.
AC013254 GI:7259653
AC013254.6 GI:7259653
HTG; HTGS_PHASE1.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 189893)
REFERENCE
AUTHORS
Celisner,S.E., Agapayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champagne,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacled,J.M., Park,S., Pfeiffer,B.,
Richardson,S., Sethi,H., Swirskas,R.R., Wan,K.H., Webster,D.,
Woollay,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
Sequencing of Drosophila melanogaster
Unpublished
2 (bases 1 to 189893)
TITLE
JOURNAL
AUTHORS
Celisner,S.E., Agapayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champagne,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacled,J.M., Park,S., Pfeiffer,B.,
Richardson,S., Sethi,H., Swirskas,R.R., Wan,K.H., Webster,D.,
Woollay,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
Direct Submission
Submitted (05-NOV-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Mar 17, 2000 this sequence version replaced gi:6984327.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 184 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 588: contig of 588 bp in length
* 589 588: gap of unknown length
* 669 1063: contig of 395 bp in length
* 1064 1143: gap of unknown length
```


Query Match	34.6%	Score 18;	DB 2;	Length 18993;
Best Local Similarity	100.0%	Pred. No. 36;		
Matches 18;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	21	CATCGCTCTGCTGCTGC	38	
Db	76500	CATCGCTCTGCTGCTGC	76483	
RESULT 38				
AC007415				
LOCUS				
DEFINITION	AC007415	198282 bp	DNA	linear HTG 17-MAR-2000
	Drosophila melanogaster chromosome 2 clone BACR05A24 (B603) RFL1-98			
	05.A.24 map 41C-41D strain Y; cn bw sp.*** SEQUENCING IN PROGRESS			
	***, 103 unordered pieces.			
AC007415				
AC007415	AC007415.11	GI:7259656		
HTG: HTGS PHASE1.				
Drosophila melanogaster (fruit fly)				
Drosophila melanogaster				
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
Ephydroidea; Drosophilidae; Drosophila.				
1 (bases 1 to 198282)				
Celniker,S.E., Agbayani,A., Arcina,T.T., Baxter,E., Blazej,R.G.,				
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,				
Doyle,C.M., Fafian,D.E., Galle,R., George,R.A., Harris,N.L.,				
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Katta,K.,				
Kearney,L., Lee,B., Lewis,S., Li,P., Lomocan,M.A., Mazda,P.,				
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacלב,J.M., Park,S., Pfeiffer,B.,				
Richards,S., Sethi,H., Svirskae,R.R., Wan,K.H., Webster,D.,				
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.				
Sequencing of Drosophila melanogaster				
Unpublished				
2 (bases 1 to 198282)				
Celniker,S.E., Agbayani,A., Arcina,T.T., Baxter,E., Blazej,R.G.,				
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,				
Doyle,C.M., Fafian,D.E., Galle,R., George,R.A., Harris,N.L.,				
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Katta,K., Kearney,L.,				
Kim,E., Lee,B., Lewis,S., Li,P., Lomocan,M.A., Mazda,P.,				
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacלב,J.M., Park,S.,				
Pfeiffer,B., Poon,L., Segre,A., Sethi,H., Snir,E.,				
Svirskae,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and				
Rubin,G.M.				
Direct Submission				
Submitted (28-APR-1999) Drosophila Genome Center, Lawrence Berkeley				
Laboratory, MS 64-121, Berkeley, CA 94720, USA				
On Mar 17, 2000 this sequence version replaced gi:7159347.				
For further information about this sequence, including its location				
and relationship to other sequences, please visit our sequence				
archive Web site (http://www.fruitfly.org/sequence/) or send email				
to baggett@fruitfly.berkeley.edu . All contigs in this submission meet				
the following cutoffs: length >= 200 bases.				
* NOTE: This is a 'working draft' sequence. It currently				
* consists of 103 contigs. The true order of the pieces				
* is not known and their order in this sequence record is				
* arbitrary. Gaps between the contigs are represented as				
* runs of N, but the exact sizes of the gaps are unknown.				
* This record will be updated with the finished sequence				
* as soon as it is available and the accession number will				
* be preserved.				
1	926:	contig of 926 bp in length		
927	1006:	gap of unknown length		
1007	1301:	contig of 295 bp in length		
1302	1381:	gap of unknown length		
1382	1860:	contig of 479 bp in length		
1861	1940:	gap of unknown length		
1941	2403:	contig of 463 bp in length		
2404	2483:	gap of unknown length		
2484	2895:	contig of 412 bp in length		
2896	2975:	gap of unknown length		

2976 3478: contig of 503 bp in length
3479 3558: gap of unknown length
3559 4034: contig of 476 bp in length
4035 4114: gap of unknown length
4115 4788: contig of 674 bp in length
4789 4868: gap of unknown length
4869 5549: contig of 681 bp in length
5550 5629: gap of unknown length
5630 6443: contig of 814 bp in length
6444 6523: gap of unknown length
6524 7697: contig of 1174 bp in length
7698 7777: gap of unknown length
7778 9243: contig of 1466 bp in length
9244 9323: gap of unknown length
9324 12162: contig of 2839 bp in length
12163 12242: gap of unknown length
12243 14263: contig of 2021 bp in length
14264 14343: gap of unknown length
14344 16104: contig of 1761 bp in length
16105 16184: gap of unknown length
16185 18228: contig of 2044 bp in length
18229 18308: gap of unknown length
18309 20874: contig of 2566 bp in length
20875 20954: gap of unknown length
20955 24951: contig of 3997 bp in length
24952 25031: gap of unknown length
25032 28493: contig of 3462 bp in length
28494 28573: gap of unknown length
28574 31026: contig of 2453 bp in length
31027 31106: gap of unknown length
31107 35546: contig of 4440 bp in length
35547 35626: gap of unknown length
35627 40849: contig of 5223 bp in length
40850 40929: gap of unknown length
40930 48761: contig of 7832 bp in length
48762 48841: gap of unknown length
48842 52656: contig of 3815 bp in length
52657 52736: gap of unknown length
52737 58188: contig of 5452 bp in length
58189 58268: gap of unknown length
58269 62896: contig of 4628 bp in length
62897 62976: gap of unknown length
62977 72175: contig of 9199 bp in length
72176 72255: gap of unknown length
72256 86109: contig of 13854 bp in length
86110 86189: gap of unknown length
86190 94772: contig of 8583 bp in length
94773 94852: gap of unknown length
94853 112088: contig of 17236 bp in length
112089 112168: gap of unknown length
112169 128075: contig of 15907 bp in length
128076 128155: gap of unknown length
128156 146641: contig of 18486 bp in length
146642 146721: gap of unknown length
146722 147805: contig of 1084 bp in length
147806 147885: gap of unknown length
147886 148190: contig of 305 bp in length
148191 148270: gap of unknown length
148271 149254: contig of 984 bp in length
149255 149334: gap of unknown length
149335 150046: contig of 712 bp in length
150047 150126: gap of unknown length
150127 151028: contig of 902 bp in length
151029 151108: gap of unknown length
151109 151716: contig of 608 bp in length
151717 151796: gap of unknown length
151797 152455: contig of 659 bp in length
152456 152535: gap of unknown length
152536 153204: contig of 669 bp in length
153205 153284: gap of unknown length
153285 153938: contig of 654 bp in length
153939 154018: gap of unknown length
154019 154280: contig of 262 bp in length

154281 154360: gap of unknown length
154361 155656: contig of 1296 bp in length
155657 155736: gap of unknown length
155737 156329: contig of 592 bp in length
156329 156408: gap of unknown length
156409 157266: contig of 858 bp in length
157267 157346: gap of unknown length
157347 157827: contig of 481 bp in length
157828 157907: gap of unknown length
157908 157928: contig of 1066 bp in length
157929 158994: gap of unknown length
158994 159073: gap of unknown length
159074 159316: contig of 243 bp in length
159317 159396: gap of unknown length
159397 159756: contig of 360 bp in length
159757 159836: gap of unknown length
159837 160697: contig of 861 bp in length
160698 160777: gap of unknown length
160778 161496: contig of 719 bp in length
161497 161576: gap of unknown length
161577 162155: contig of 579 bp in length
162156 162236: gap of unknown length
162236 162542: contig of 307 bp in length
162543 162682: gap of unknown length
162682 162964: contig of 342 bp in length
162965 163044: gap of unknown length
163045 163544: contig of 500 bp in length
163545 163624: gap of unknown length
163625 165092: contig of 1468 bp in length
165093 165172: gap of unknown length
165173 165866: contig of 694 bp in length
165867 165946: gap of unknown length
165947 166183: contig of 237 bp in length
166184 166263: gap of unknown length
166264 167264: contig of 1000 bp in length
167264 167343: gap of unknown length
167344 167883: contig of 540 bp in length
167884 167963: gap of unknown length
167964 168338: contig of 375 bp in length
168339 168418: gap of unknown length
168419 168719: contig of 301 bp in length
168720 168799: gap of unknown length
168800 169480: contig of 681 bp in length
169481 169560: gap of unknown length
169561 170067: contig of 507 bp in length
170068 170147: gap of unknown length
170148 170904: contig of 757 bp in length
170905 170984: gap of unknown length
170985 171769: contig of 785 bp in length
171770 171849: gap of unknown length
171849 172398: contig of 549 bp in length
172399 172478: gap of unknown length
172479 173200: contig of 722 bp in length
173201 173280: gap of unknown length
173281 173833: contig of 553 bp in length
173834 173913: gap of unknown length
173914 174940: contig of 1027 bp in length
174941 175020: gap of unknown length
175021 175732: contig of 712 bp in length
175733 175813: gap of unknown length
175813 176895: contig of 1083 bp in length
176896 176975: gap of unknown length
176976 177582: contig of 607 bp in length
177583 177662: gap of unknown length
177663 178191: contig of 449 bp in length
178192 178900: contig of 709 bp in length
178901 178980: gap of unknown length

Query Match 34.6%; Score 18; DB 2; Length 198282;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 CATGCTCTGCTGCTGC 38

Db 93843 CATCGCTCTCGTCGCTGC 93860

RESULT 39	LOCUS	DEFINITION	DNA	HTG
AC007415/c	AC007415	Drosophila melanogaster chromosome 2 clone BACR05.224 (D603)	198282 bp	1 linear
				RPT1-99

ACCESSION AC007415
VERSION AC007415.11 GI:7259656
KEYWORDS HTG; HTGS_PPhase1.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster

REFERENCE	AUTHORS
1 (bases 1 to 198282)	Celniker, S.E., Agbavanyi, A., Arcaina, T.T., Baxter, E., Blazel, R.G.

TITLE	JOURNAL	REFERENCE	AUTHORS
Sequencing of <i>Drosophila melanogaster</i> Unpublished		2 (bases 1 to 198282)	Ceolnter, S.E., Agbayan, A., Arcaina, T.T., Baxter, E., Blazet, R.G.

Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacled, J.M., Park, S., Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Shtyr, E., Svayaskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and Rudin, G.M.

TITLE	JOURNAL	COMMENT
Direct Submission	Submitted (28-APR-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA	<p>On Mar 17, 2000 this sequence version replaced gi:7159347. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (http://www.fruitfly.org/sequence/) or send email to hdpp@fruitfly.berkeley.edu. All contigs in this submission meet the following cutoffs: length >= 200 bases.</p> <p>NOTE: This is a 'working draft' sequence. It currently consists of 103 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.</p>

*	1	926:	contig of 926 bp in length
*	927	1006:	gap of unknown length
*	1007	1301:	contig of 285 bp in length
*	1302	1381:	gap of unknown length
*	1382	1860:	contig of 479 bp in length
*	1861	1940:	gap of unknown length
*	1941	2403:	contig of 463 bp in length
*	2404	2483:	gap of unknown length
*	2484	2895:	contig of 412 bp in length
*	2896	2975:	gap of unknown length
*	2976	3478:	contig of 503 bp in length
*	3479	3558:	gap of unknown length
*	3559	4034:	contig of 476 bp in length
*	4035	4114:	gap of unknown length
*	4115	4788:	contig of 674 bp in length
*	4789	4868:	gap of unknown length
*	4869	5549:	contig of 681 bp in length

550	5629:	gap of unknown length
5630	6443:	contig of 814 bp in length
5644	6523:	gap of unknown length
6524	7697:	contig of 1174 bp in length
7658	7777:	gap of unknown length
7778	9243:	contig of 1466 bp in length
9244	9323:	gap of unknown length
9324	12162:	contig of 2839 bp in length
12163	12242:	gap of unknown length
12243	14263:	contig of 2021 bp in length
14264	14343:	gap of unknown length
14344	16104:	contig of 1761 bp in length
16105	16184:	gap of unknown length
16185	18228:	contig of 2044 bp in length
18229	18308:	gap of unknown length
18309	20874:	contig of 2566 bp in length
20875	20954:	gap of unknown length
20955	24951:	contig of 3997 bp in length
24952	25031:	gap of unknown length
25032	28493:	contig of 3462 bp in length
28494	28573:	gap of unknown length
28574	31026:	contig of 2453 bp in length
31027	31106:	gap of unknown length
31107	35546:	contig of 4440 bp in length
35447	35626:	gap of unknown length
35627	40849:	contig of 5223 bp in length
40850	40929:	gap of unknown length
40930	48761:	contig of 7832 bp in length
48762	48841:	gap of unknown length
48842	52656:	contig of 3815 bp in length
52657	55736:	gap of unknown length
55737	58188:	contig of 5452 bp in length
58189	62896:	gap of unknown length
62897	62976:	gap of unknown length
62977	72175:	contig of 919 bp in length
72176	72255:	gap of unknown length
72256	86109:	contig of 13854 bp in length
86110	86189:	gap of unknown length
86130	94772:	contig of 853 bp in length
94773	94852:	gap of unknown length
94853	112088:	contig of 1736 bp in length
112089	112168:	gap of unknown length
112169	128075:	contig of 15007 bp in length
128076	128155:	gap of unknown length
128156	146641:	contig of 18466 bp in length
146642	146721:	gap of unknown length
146722	147805:	contig of 1084 bp in length
147806	147885:	gap of unknown length
147886	148190:	contig of 305 bp in length
148191	148270:	gap of unknown length
148271	149254:	contig of 984 bp in length
149255	149334:	gap of unknown length
149335	150046:	contig of 712 bp in length
150047	150126:	gap of unknown length
150127	151028:	contig of 902 bp in length
151029	151108:	gap of unknown length
151109	151716:	contig of 608 bp in length
151717	151796:	gap of unknown length
151797	152455:	contig of 659 bp in length
152456	152535:	gap of unknown length
152536	153204:	contig of 669 bp in length
153205	153288:	gap of unknown length
153289	153938:	contig of 654 bp in length
153939	154018:	gap of unknown length
154019	154280:	contig of 262 bp in length
154281	154360:	gap of unknown length
154361	155656:	contig of 1296 bp in length
155657	155736:	gap of unknown length
155737	156128:	contig of 592 bp in length
156129	156408:	gap of unknown length
156409	157266:	contig of 858 bp in length
157267	157346:	gap of unknown length

* 157347 157827: contig of 481 bp in length
 * 157828 157907: gap of unknown length
 * 157908 158993: contig of 1086 bp in length
 * 158994 159073: gap of unknown length
 * 159074 159116: contig of 243 bp in length
 * 159317 159396: gap of unknown length
 * 159397 159756: contig of 360 bp in length
 * 159757 159836: gap of unknown length
 * 159837 160697: contig of 861 bp in length
 * 160698 160777: gap of unknown length
 * 160778 161496: contig of 719 bp in length
 * 161497 161576: gap of unknown length
 * 161577 162155: contig of 579 bp in length
 * 162156 162235: gap of unknown length
 * 162236 162542: contig of 307 bp in length
 * 162543 162622: gap of unknown length
 * 162623 162964: contig of 342 bp in length
 * 162965 163044: gap of unknown length
 * 163045 163544: contig of 500 bp in length
 * 163545 163624: gap of unknown length
 * 163625 165092: contig of 1468 bp in length
 * 165093 165172: gap of unknown length
 * 165173 165866: contig of 694 bp in length
 * 165867 165946: gap of unknown length
 * 165947 166183: contig of 237 bp in length
 * 166184 166263: gap of unknown length
 * 166264 167264: contig of 1000 bp in length
 * 167264 167343: gap of unknown length
 * 167344 167883: contig of 540 bp in length
 * 167884 167963: gap of unknown length
 * 167964 168338: contig of 375 bp in length
 * 168339 168418: gap of unknown length
 * 168419 168719: contig of 301 bp in length
 * 168720 168799: gap of unknown length
 * 168800 169480: contig of 681 bp in length
 * 169481 169560: gap of unknown length
 * 169561 170057: contig of 507 bp in length
 * 170058 170147: gap of unknown length
 * 170148 170904: contig of 757 bp in length
 * 170905 170984: gap of unknown length
 * 170985 171769: contig of 785 bp in length
 * 171770 171849: gap of unknown length
 * 171850 172398: contig of 549 bp in length
 * 172399 172478: gap of unknown length
 * 172479 173200: contig of 722 bp in length
 * 173201 173280: gap of unknown length
 * 173281 173833: contig of 553 bp in length
 * 173834 173913: gap of unknown length
 * 173914 174940: contig of 1027 bp in length
 * 174941 175020: gap of unknown length
 * 175021 175732: contig of 712 bp in length
 * 175733 175812: gap of unknown length
 * 175813 176895: contig of 1083 bp in length
 * 176896 176975: gap of unknown length
 * 176976 177582: contig of 607 bp in length
 * 177583 177662: gap of unknown length
 * 177663 178111: contig of 449 bp in length
 * 178112 178191: gap of unknown length
 * 178192 178900: contig of 709 bp in length
 * 178901 178980: gap of unknown length

Query Match 34.6%; Score 18; DB 2; Length 198282;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 21 CATCGCTCTGCTGCTGC 38
 Db 146075 CATCGCTCTGCTGCTGC 146058

RESULT 40
 AC013831 AC013831 201451 bp DNA linear HTG 15-MAR-2002
 LOCUS

DEFINITION Drosophila melanogaster chromosome 2 clone BAC12A16 (D1172)
 ACCESSION RPCI-98 12.A.16 map 40D-40D strain y: cn bw sp, *** SEQUENCING IN
 PROGRESS ***, 136 unordered pieces.
 AC013831
 VERSION AC013831.13 GI:19482321
 KEYWORDS HTG, HTGS, PHASE1
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 201451)
 AUTHORS Celniker, S.E., Abmayyan, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
 Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
 Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
 Hinkle, A., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K.,
 Kearney, L., Lee, B., Lewis, S., Li, P., Ling, H., Moshrefi, A.R.,
 Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B.,
 Richards, S., Sethi, H., Svirskas, R.R., Wan, K.H., Webster, D.,
 Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G.M.
 TITLE Direct Submission
 JOURNAL Submitted (15-NOV-1999) Drosophila Genome Center, Lawrence Berkeley
 Laboratory, MS 64-121, Berkeley, CA 94720, USA
 REFERENCE 3 (bases 1 to 201451)
 AUTHORS Celniker, S.E., Abmayyan, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
 Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
 Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
 Hinkle, A., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K.,
 Kearney, L., Lee, B., Lewis, S., Li, P., Ling, H., Moshrefi, A.R.,
 Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B.,
 Richards, S., Sethi, H., Svirskas, R.R., Wan, K.H., Webster, D.,
 Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G.M.
 TITLE Direct Submission
 JOURNAL Submitted (15-MAR-2002) Drosophila Genome Center, Lawrence Berkeley
 Laboratory, MS 64-121, Berkeley, CA 94720, USA
 COMMENT For further information about this sequence, including its location
 and relationship to other sequences, please visit our sequence
 archive Web site (<http://www.fruitfly.org/sequence/>) or send email
 to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
 the following cutoffs: length >= 200 bases.
 NOTE: This is a 'working draft' sequence. It currently
 consists of 136 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1 622: contig of 622 bp in length
 * 623 702: gap of unknown length
 * 703 1319: contig of 617 bp in length
 * 1320 1399: gap of unknown length
 * 1400 2018: contig of 619 bp in length
 * 2019 2098: gap of unknown length
 * 2099 2696: contig of 598 bp in length
 * 2697 2776: gap of unknown length
 * 2777 3421: contig of 645 bp in length
 * 3422 4078: gap of unknown length
 * 4079 4158: contig of 577 bp in length
 * 4159 4671: contig of 513 bp in length

* 4672 4751: gap of unknown length
* 4752 5517: contig of 766 bp in length
* 5518 5597: gap of unknown length
* 5598 6243: contig of 646 bp in length
* 6244 6323: gap of unknown length
* 6324 6965: contig of 642 bp in length
* 6966 7045: gap of unknown length
* 7046 7659: contig of 614 bp in length
* 7660 7739: gap of unknown length
* 7740 8525: contig of 786 bp in length
* 8526 8605: gap of unknown length
* 8606 9300: contig of 695 bp in length
* 9301 9380: gap of unknown length
* 9381 10413: contig of 1033 bp in length
* 10414 10493: gap of unknown length
* 10494 11059: contig of 566 bp in length
* 11060 11139: gap of unknown length
* 11140 12030: contig of 891 bp in length
* 12031 12110: gap of unknown length
* 12111 12820: contig of 710 bp in length
* 12821 12900: gap of unknown length
* 12901 13575: contig of 675 bp in length
* 13576 13655: gap of unknown length
* 13656 14388: contig of 733 bp in length
* 14389 14468: gap of unknown length
* 14469 15443: contig of 975 bp in length
* 15444 15523: gap of unknown length
* 15524 16504: contig of 981 bp in length
* 16505 16584: gap of unknown length
* 16585 18119: contig of 1535 bp in length
* 18120 18199: gap of unknown length
* 18200 19453: contig of 1254 bp in length
* 19454 19533: gap of unknown length
* 19534 20795: contig of 1262 bp in length
* 20796 20875: gap of unknown length
* 20876 22136: contig of 1261 bp in length
* 22137 22216: gap of unknown length
* 22217 23747: contig of 1531 bp in length
* 23748 23827: gap of unknown length
* 23828 25601: contig of 1774 bp in length
* 25602 25681: gap of unknown length
* 25682 28215: contig of 2534 bp in length
* 28216 28295: gap of unknown length
* 28296 29804: contig of 1509 bp in length
* 29805 29884: gap of unknown length
* 29885 31373: contig of 1489 bp in length
* 31374 31453: gap of unknown length
* 31454 33841: contig of 2388 bp in length
* 33842 33921: gap of unknown length
* 33922 37415: contig of 3494 bp in length
* 37416 37495: gap of unknown length
* 37496 40096: contig of 2601 bp in length
* 40097 40176: gap of unknown length
* 40177 44443: contig of 4267 bp in length
* 44444 44523: gap of unknown length
* 44524 47877: contig of 3354 bp in length
* 47878 47957: gap of unknown length
* 47958 52119: contig of 4162 bp in length
* 52120 52199: gap of unknown length
* 52200 56290: contig of 4091 bp in length
* 56291 56370: gap of unknown length
* 56371 62684: contig of 6314 bp in length
* 62685 62764: gap of unknown length
* 62765 71961: contig of 9197 bp in length
* 71962 72041: gap of unknown length
* 72042 79797: contig of 7756 bp in length
* 79798 79877: gap of unknown length
* 79878 101310: contig of 21433 bp in length
* 101311 101390: gap of unknown length
* 101391 109584: contig of 8194 bp in length
* 109585 109664: gap of unknown length
* 109665 121016: contig of 13352 bp in length
* 121017 121096: gap of unknown length

* 121097 137479: contig of 16383 bp in length
* 137480 137559: gap of unknown length
* 137560 138592: contig of 1033 bp in length
* 138593 138672: gap of unknown length
* 138673 139477: contig of 805 bp in length
* 139478 139557: gap of unknown length
* 139558 140520: contig of 963 bp in length
* 140521 140600: gap of unknown length
* 140601 141218: contig of 618 bp in length
* 141219 141298: gap of unknown length
* 141299 141969: contig of 671 bp in length
* 141970 142049: gap of unknown length
* 142050 142463: contig of 414 bp in length
* 142464 142543: gap of unknown length
* 142544 143220: contig of 677 bp in length
* 143221 143300: gap of unknown length
* 143301 143706: contig of 406 bp in length
* 143707 143786: gap of unknown length
* 143787 144114: contig of 328 bp in length
* 144115 144194: gap of unknown length
* 144195 144798: contig of 604 bp in length
* 144799 144878: gap of unknown length
* 144879 145387: contig of 509 bp in length
* 145388 145467: gap of unknown length
* 145468 146605: contig of 1138 bp in length
* 146606 146685: gap of unknown length
* 146886 146934: contig of 249 bp in length
* 146935 147014: gap of unknown length
* 147015 147251: contig of 237 bp in length
* 147252 147331: gap of unknown length
* 147332 148015: contig of 684 bp in length
* 148016 148095: gap of unknown length
* 148096 148714: contig of 619 bp in length
* 148715 148794: gap of unknown length
* 148795 149416: contig of 622 bp in length
* 149417 149496: gap of unknown length
* 149497 150142: contig of 646 bp in length
* 150143 150222: gap of unknown length
* 150223 150916: contig of 694 bp in length
* 150917 150996: gap of unknown length
* 150997 151559: contig of 563 bp in length
* 151560 151639: gap of unknown length
* 151640 152228: contig of 589 bp in length
* 152229 152308: gap of unknown length
* 152309 152918: contig of 610 bp in length
* 152919 152998: gap of unknown length
* 152999 153639: contig of 641 bp in length
* 153640 153719: gap of unknown length
* 153720 154031: contig of 312 bp in length
* 154032 154111: gap of unknown length
* 154112 154581: contig of 470 bp in length
* 154582 154661: gap of unknown length
* 154662 154873: contig of 212 bp in length

Query Match 34.6% Score 18; DB 2; Length 201451;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 CATCGCTCTCGTCGCTGC 38
DB 106212 CATCGCTCTCGTCGCTGC 106229

Search completed: November 13, 2003, 10:26:37
Job time : 2042 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2003, 08:53:26 Search time 217 Seconds
(without alignments)
646.870 Million cell updates/sec

Title: US-10-081-935-1

Perfect score: 52
Sequence: 1 atgcagctctctccacgctct.....cgctgcgcgcctcgcacgctg 52

Scoring table: OLIGO NUC
Gapop_60.0, Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

N_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	19	36.5	3416	21	AA552247	Maize glycine-rich
C 2	17	32.7	1392	25	AA050552	Mycobacterium aviu
C 3	17	32.7	2684	23	ABU19972	Drosophila melanog
C 4	17	32.7	4310	23	ABU15450	Drosophila melanog
5	16	30.8	276	21	AAA31959	Plant microsatelli
6	16	30.8	357	21	AAA31675	Plant microsatelli
7	16	30.8	438	21	AAA31705	Plant microsatelli
C 8	16	30.8	598	23	AA584162	DNA encoding novel

C 9	16	30.8	638	21	AA035333	Arabidopsis thalia
10	16	30.8	1582	23	AA569014	DNA encoding novel
11	16	30.8	1653	24	AA028124	Human leucine-rich
12	16	30.8	1872	22	AA076853	Human secreted pro
13	16	30.8	2369	24	AA038695	Human Lp220 secret
14	16	30.8	2451	24	ABV99421	Human NOV44b codin
15	16	30.8	2563	24	ABV99422	Human NOV44c codin
16	16	30.8	2569	25	AAK99609	MDMT related human
17	16	30.8	3201	24	ABV99420	Human NOV44a codin
C 18	16	30.8	3947	21	AA016203	Human prostate can
C 19	16	30.8	3953	22	AAH34839	Human colon cancer
C 20	16	30.8	6507	23	AAH15493	Drosophila melanog
C 21	16	30.8	7434	23	AA584163	DNA encoding novel
C 22	16	30.8	15435	23	ABU14186	Drosophila melanog
C 23	16	30.8	47999	25	AA052898	Human tweety homol
C 24	16	30.8	59667	23	ABU15492	Drosophila melanog
C 25	16	30.8	349980	24	ABO81844	Bifidobacterium lo
C 26	15	28.8	214	24	ABL74833	Corn tassell-derive
C 27	15	28.8	264	25	ABX33704	Human GDP-mannose
C 28	15	28.8	335	25	ABX17641	S2 subtracction lib
C 29	15	28.8	336	25	ABX19200	Human GDP-mannose
C 30	15	28.8	357	22	AA531315	Human cDNA encodin
C 31	15	28.8	357	23	ABK3452	DNA encoding novel
C 32	15	28.8	357	23	ABQ66639	Human polynucleoti
C 33	15	28.8	416	23	AA576408	DNA encoding novel
C 34	15	28.8	423	20	AA219396	M. tuberculosis an
C 35	15	28.8	423	20	AA219396	M. tuberculosis re
C 36	15	28.8	448	21	ABQ62561	Mycobacterium tube
C 37	15	28.8	459	22	AA190273	Human secreted pro
C 38	15	28.8	471	21	AA020001	Human foetal liver
C 39	15	28.8	481	22	ABA58638	Probe #6099 for ge
C 40	15	28.8	481	22	ABA27633	Human brain expres
C 41	15	28.8	481	22	AAK06770	Human bone marrow
C 42	15	28.8	481	22	AAK24463	Human genome-deriv
C 43	15	28.8	481	22	AA116089	Human immune-haema
C 44	15	28.8	481	22	AA116089	Human cDNA clone (
C 45	15	28.8	481	22	AA138312	Aspergillus oryzae
C 46	15	28.8	481	22	AB532191	Human cDNA clone (
C 47	15	28.8	481	23	AB507268	M. capuulatus gene
C 48	15	28.8	619	22	AAK06662	Human genome-deriv
C 49	15	28.8	667	22	AAH05960	Human immune-haema
C 50	15	28.8	695	21	AA012705	Human cDNA clone (
C 51	15	28.8	753	22	AAH08589	Aspergillus oryzae
C 52	15	28.8	758	24	ABO90246	Human cDNA clone (
C 53	15	28.8	1197	22	AA158291	M. capuulatus gene
C 54	15	28.8	1340	21	AA000690	Human polynucleoti
C 55	15	28.8	1358	22	AA541256	Human hydroxylase pr
C 56	15	28.8	1366	23	ABV23063	cDNA encoding nove
C 57	15	28.8	1386	23	ABV28899	Human prostate exp
C 58	15	28.8	1404	25	ACC50288	Human prostate exp
C 59	15	28.8	1431	24	ABT07729	Breast cancer asso
C 60	15	28.8	1435	25	ABX72219	Breast cancer-asso
C 61	15	28.8	1446	22	AAH46450	Human NOVX polynuc
C 62	15	28.8	1473	24	ABQ60922	Penicillium chrysos
C 63	15	28.8	1502	21	AA033695	Human hippocampal
C 64	15	28.8	1507	21	AA033595	Arabidopsis thalia
C 65	15	28.8	1554	25	ABT20720	Arabidopsis thalia
C 66	15	28.8	1750	24	AB567741	Aspergillus fumiga
C 67	15	28.8	1778	25	ABT20122	Human intracellular
C 68	15	28.8	1878	23	AAH16752	Aspergillus fumiga
C 69	15	28.8	2000	24	ABE120539	Drosophila melanog
C 70	15	28.8	2134	23	ABE15173	Arabidopsis thalia
C 71	15	28.8	2304	23	ABU02831	Drosophila melanog
C 72	15	28.8	2454	23	ABU17476	Drosophila melanog
C 73	15	28.8	2520	24	AA551556	Pseudomonas aerugi
C 74	15	28.8	2556	22	AAH17696	Human HCCA2 encodi
C 75	15	28.8	2617	22	AAH16752	Human cDNA sequenc
C 76	15	28.8	2672	21	AAH16680	Human secreted pro
C 77	15	28.8	2702	24	ABA99189	Human HCCA1 encodi
C 78	15	28.8	2790	24	ABU65208	Lung cancer relat
C 79	15	28.8	2790	24	ABU67590	Ovary cancer relat
C 80	15	28.8	3103	23	ABU08033	Drosophila melanog
C 81	15	28.8	3140	25	ABT17712	Aspergillus fumiga
81	15	28.8	3323	24	ABK35806	cDNA sequence #197

C 82	15	28.8	3465	21	AAAT5900	DNA encoding a bio
C 83	15	28.8	3778	25	ABT19526	Aspergillus fumiga
C 84	15	28.8	3830	25	ABL19782	Drosophila melanog
C 85	15	28.8	3878	24	ABL20538	Drosophila melanog
C 86	15	28.8	3963	24	ABA90499	Drosophila cell cy
C 87	15	28.8	4198	22	AAA89296	Arabidopsis enoyl-
C 88	15	28.8	4338	12	AAO11826	Encodes N-terminal
C 89	15	28.8	4382	23	ABL1565	Drosophila melanog
C 90	15	28.8	4649	10	AAV90659	Sequence encoding
C 91	15	28.8	4763	21	AACT6534	Human ORFX ORF2089
C 92	15	28.8	5277	23	ABL08995	Drosophila melanog
C 93	15	28.8	5676	23	ABL02830	Drosophila melanog
C 94	15	28.8	5721	11	AAO06613	Ageryl cyclase gen
C 95	15	28.8	6059	24	ABL54343	Chemically treated
C 96	15	28.8	6201	23	ABL08032	Drosophila melanog
C 97	15	28.8	6381	10	AAV91167	Sequence encoding
C 98	15	28.8	6442	24	ABK50870	Bordetella pertuss
C 99	15	28.8	7382	23	ABL13554	Drosophila melanog
C 100	15	28.8	7382	23	ABL13564	Drosophila melanog
C 101	15	28.8	9405	24	ABO93392	Human cDNA SEQ ID
C 102	15	28.8	9926	25	ABZ71998	Human ASH1 cDNA Ge
C 103	15	28.8	10676	23	ABV23147	Human prostate exp
C 104	15	28.8	10676	23	ABV28988	Human prostate exp
C 105	15	28.8	10715	23	ABL08994	Drosophila melanog
C 106	15	28.8	11847	22	AA557196	DNA encoding Droso
C 107	15	28.8	11847	23	ABL29442	Drosophila melanog
C 108	15	28.8	13545	22	ABA21471	Human nervous syst
C 109	15	28.8	13545	22	AAK80349	Human immune/haema
C 110	15	28.8	13676	22	AAK80349	Human immune/haema
C 111	15	28.8	38110	25	AAI53730	Ribonuclease P RNA
C 112	15	28.8	44377	18	AAAT78508	Platenolide syntha
C 113	15	28.8	44377	18	AAAT80414	Human immune/haema
C 114	15	28.8	48561	22	AAK82012	Sorngium celluloso
C 115	15	28.8	65750	21	AAZ55887	Sorngium celluloso
C 116	15	28.8	71989	21	AAA29349	Sorngium celluloso
C 117	15	28.8	90600	24	ABO78872	S. roseosporus dap
C 118	15	28.8	349980	24	ABO81845	Bifidobacterium lo
C 119	15	28.8	349980	24	ABO81845	Bifidobacterium lo
C 120	15	28.8	349980	24	ABO81845	Bifidobacterium lo
C 121	15	28.8	534720	19	AAV30458	Rhizobium species
C 122	15	28.8	536165	19	AAV30459	Rhizobium species
C 123	15	28.8	4403765	22	AAI99683	Mycobacterium tube
C 124	15	28.8	4403765	22	AAI99683	Mycobacterium tube
C 125	15	28.8	4411529	22	AAI99682	Mycobacterium tube
C 126	15	28.8	4411529	22	AAI99682	Mycobacterium tube
C 127	14	26.9	20	22	AAH63071	Shrimp white spot
C 128	14	26.9	135	21	AAAC04546	Human secreted pro
C 129	14	26.9	208	21	AAAC08471	Human secreted pro
C 130	14	26.9	234	24	ABT06874	Human ovarian carc
C 131	14	26.9	234	25	ABX72752	Human ovarian carc
C 132	14	26.9	266	24	ABN75313	Human ORF260 cDNA,
C 133	14	26.9	280	21	AAA37428	Maize acyltransfer
C 134	14	26.9	284	21	AAA37427	Maize acyltransfer
C 135	14	26.9	287	21	AAA37425	Maize acyltransfer
C 136	14	26.9	291	21	AAA37425	Maize acyltransfer
C 137	14	26.9	300	21	AAA37426	Maize acyltransfer
C 138	14	26.9	308	21	AAZ61715	cDNA encoding murti
C 139	14	26.9	308	22	AAAC9648	Skin cell cDNA, SE
C 140	14	26.9	310	22	ABL34800	Murine cDNA isolat
C 141	14	26.9	314	21	AAA37424	Maize acyltransfer
C 142	14	26.9	345	21	AAAC07690	Human secreted pro
C 143	14	26.9	368	21	AAAC56456	Human secreted pro
C 144	14	26.9	368	21	AAAC57136	Pinus radiata tran
C 145	14	26.9	370	15	AAO55250	Pinus radiata tran
C 146	14	26.9	384	25	ABX19448	Grapevine ribosoma
C 147	14	26.9	390	22	AAE64610	Human GDP-mannose
C 148	14	26.9	399	22	AAI82496	Novel human polyu
C 149	14	26.9	427	21	AAA65872	Human polynucleoti
C 150	14	26.9	438	21	AAAC5711	E. coli proliferat
C 151	14	26.9	446	24	ABN55029	Zea mays DNA fragm
C 152	14	26.9	474	21	AAA52709	Human ORFX polynuc
C 153	14	26.9	513	21	AAAC41052	Wheat Mt1 homology
C 154	14	26.9	515	24	AAI46712	Zea mays DNA fragm
C 155	14	26.9	522	22	ABA19450	Human nervous syst
C 156	14	26.9	522	22	ABA19451	Human nervous syst
C 157	14	26.9	522	22	AAK83063	Human immune/haema
C 158	14	26.9	522	22	AAK83064	Human immune/haema
C 159	14	26.9	522	22	AAK83064	Human immune/haema
C 160	14	26.9	522	22	AAK83064	Human immune/haema
C 161	14	26.9	522	22	AAK83064	Human immune/haema
C 162	14	26.9	522	22	AAK83064	Human immune/haema
C 163	14	26.9	522	22	AAK83064	Human immune/haema
C 164	14	26.9	522	22	AAK83064	Human immune/haema
C 165	14	26.9	522	22	AAK83064	Human immune/haema
C 166	14	26.9	522	22	AAK83064	Human immune/haema
C 167	14	26.9	522	22	AAK83064	Human immune/haema
C 168	14	26.9	522	22	AAK83064	Human immune/haema
C 169	14	26.9	522	22	AAK83064	Human immune/haema
C 170	14	26.9	522	22	AAK83064	Human immune/haema
C 171	14	26.9	522	22	AAK83064	Human immune/haema
C 172	14	26.9	522	22	AAK83064	Human immune/haema
C 173	14	26.9	522	22	AAK83064	Human immune/haema
C 174	14	26.9	522	22	AAK83064	Human immune/haema
C 175	14	26.9	522	22	AAK83064	Human immune/haema
C 176	14	26.9	522	22	AAK83064	Human immune/haema
C 177	14	26.9	522	22	AAK83064	Human immune/haema
C 178	14	26.9	522	22	AAK83064	Human immune/haema
C 179	14	26.9	522	22	AAK83064	Human immune/haema
C 180	14	26.9	522	22	AAK83064	Human immune/haema
C 181	14	26.9	522	22	AAK83064	Human immune/haema
C 182	14	26.9	522	22	AAK83064	Human immune/haema
C 183	14	26.9	522	22	AAK83064	Human immune/haema
C 184	14	26.9	522	22	AAK83064	Human immune/haema
C 185	14	26.9	522	22	AAK83064	Human immune/haema
C 186	14	26.9	522	22	AAK83064	Human immune/haema
C 187	14	26.9	522	22	AAK83064	Human immune/haema
C 188	14	26.9	522	22	AAK83064	Human immune/haema
C 189	14	26.9	522	22	AAK83064	Human immune/haema
C 190	14	26.9	522	22	AAK83064	Human immune/haema
C 191	14	26.9	522	22	AAK83064	Human immune/haema
C 192	14	26.9	522	22	AAK83064	Human immune/haema
C 193	14	26.9	522	22	AAK83064	Human immune/haema
C 194	14	26.9	522	22	AAK83064	Human immune/haema
C 195	14	26.9	522	22	AAK83064	Human immune/haema
C 196	14	26.9	522	22	AAK83064	Human immune/haema
C 197	14	26.9	522	22	AAK83064	Human immune/haema
C 198	14	26.9	522	22	AAK83064	Human immune/haema
C 199	14	26.9	522	22	AAK83064	Human immune/haema
C 200	14	26.9	522	22	AAK83064	Human immune/haema
C 201	14	26.9	522	22	AAK83064	Human immune/haema
C 202	14	26.9	522	22	AAK83064	Human immune/haema
C 203	14	26.9	522	22	AAK83064	Human immune/haema
C 204	14	26.9	522	22	AAK83064	Human immune/haema
C 205	14	26.9	522	22	AAK83064	Human immune/haema
C 206	14	26.9	522	22	AAK83064	Human immune/haema
C 207	14	26.9	522	22	AAK83064	Human immune/haema
C 208	14	26.9	522	22	AAK83064	Human immune/haema
C 209	14	26.9	522	22	AAK83064	Human immune/haema
C 210	14	26.9	522	22	AAK83064	Human immune/haema
C 211	14	26.9	522	22	AAK83064	Human immune/haema
C 212	14	26.9	522	22	AAK83064	Human immune/haema
C 213	14	26.9	522	22	AAK83064	Human immune/haema
C 214	14	26.9	522	22	AAK83064	Human immune/haema
C 215	14	26.9	522	22	AAK83064	Human immune/haema
C 216	14	26.9	522	22	AAK83064	Human immune/haema
C 217	14	26.9	522	22	AAK83064	Human immune/haema
C 218	14	26.9	522	22	AAK83064	Human immune/haema
C 219	14	26.9	522	22	AAK83064	Human immune/haema
C 220	14	26.9	522	22	AAK83064	Human immune/haema
C 221	14	26.9	522	22	AAK83064	Human immune/haema
C 222	14	26.9	522	22	AAK83064	Human immune/haema
C 223	14	26.9	522	22	AAK83064	Human immune/haema
C 224	14	26.9	522	22	AAK83064	Human immune/haema
C 225	14	26.9	522	22	AAK83064	Human immune/haema
C 226	14	26.9	522	22	AAK83064	Human immune/haema
C 227	14	26.9	522	22	AAK83064	Human immune/haema

C 228	14	26.9	2036	22	AAH74908	cdna sequence iso1	C 301	14	26.9	5494	23	ABL13956	Drosophila melanog
C 229	14	26.9	2085	22	AAH23690	Human tumour suppr	C 302	14	26.9	5509	24	ABA92640	Bifidobacterium bi
C 230	14	26.9	2098	22	AAS05387	Human caspase recr	C 303	14	26.9	5612	24	ABL57471	Human proteoin phos
C 231	14	26.9	2098	24	ABA00332	Human CARP-9 cdna.	C 304	14	26.9	6029	24	ABL17630	Drosophila melanog
C 232	14	26.9	2133	21	AAZ29828	Herpes simplex vir	C 305	14	26.9	6208	23	ABL32195	Human immune syste
C 233	14	26.9	2176	22	AAH18321	Human cdna sequenc	C 306	14	26.9	6236	19	AAH75759	Human neurofilamen
C 234	14	26.9	2183	23	ABX11432	DNA encoding caspa	C 307	14	26.9	6439	21	AAH75235	Nucleotide sequenc
C 235	14	26.9	2351	23	ABL28797	Drosophila melanog	C 308	14	26.9	6454	23	ABL29284	Drosophila melanog
C 236	14	26.9	2351	23	ABL19786	Drosophila melanog	C 309	14	26.9	6502	22	AAF98718	Human late stage o
C 237	14	26.9	2534	20	AAZ27285	M. prunae DNA poly	C 310	14	26.9	6513	21	AAH75234	Nucleotide sequenc
C 238	14	26.9	2534	24	AAH35584	Metalloislaera pru	C 311	14	26.9	6553	23	ABL10330	Drosophila melanog
C 239	14	26.9	2534	25	ABX14688	DNA encoding Metal	C 312	14	26.9	6625	23	ABL19750	Drosophila melanog
C 240	14	26.9	2641	23	ABL28434	Drosophila melanog	C 313	14	26.9	6840	24	ABK84742	Human cdna differe
C 241	14	26.9	2769	22	AAH62803	Shrimp white spot	C 314	14	26.9	6840	24	ABL88419	Pain regulated cDN
C 242	14	26.9	2792	23	ABL10062	Drosophila melanog	C 315	14	26.9	6843	18	AAH84509	Human LDL receptor
C 243	14	26.9	2792	24	ABK14729	Barley DEAD box pr	C 316	14	26.9	6902	22	AAH98724	Human late stage o
C 244	14	26.9	2861	18	AAH64785	Human oxygen regul	C 317	14	26.9	6906	25	ACC50992	Human late stage o
C 245	14	26.9	2894	25	ABT17875	Aspergillus fumiga	C 318	14	26.9	6981	24	ABH76520	CDNA encoding huma
C 246	14	26.9	2902	11	AAQ06633	Clone encoding rat	C 319	14	26.9	6981	24	ABK64741	Human benign prost
C 247	14	26.9	2904	23	ABL25675	Drosophila melanog	C 320	14	26.9	6981	24	ABL68557	Kidney cancer rela
C 248	14	26.9	2969	23	ABL17631	Drosophila melanog	C 321	14	26.9	7140	24	ABK51214	Rat cdna encoding
C 249	14	26.9	3053	25	ABT19689	Aspergillus fumiga	C 322	14	26.9	7140	24	ABK51223	Rat Kidins220 geno
C 250	14	26.9	3078	22	AAI93905	Human stomach can	C 323	14	26.9	7251	23	ABL21046	Drosophila melanog
C 251	14	26.9	3078	22	AAH16495	Human cdna sequenc	C 324	14	26.9	7251	24	AAH94894	Human DNA sequenc
C 252	14	26.9	3096	23	ABL13957	Drosophila melanog	C 325	14	26.9	7319	24	AAH94894	Human breast cance
C 253	14	26.9	3105	23	ABL18032	Drosophila melanog	C 326	14	26.9	7460	25	ABT31940	Drosophila cell cy
C 254	14	26.9	3150	24	ABK13117	Human secretory po	C 327	14	26.9	7494	24	ABH08774	Drosophila melanog
C 255	14	26.9	3155	23	ABL28776	Drosophila melanog	C 328	14	26.9	7920	23	ABL26882	Drosophila melanog
C 256	14	26.9	3163	24	AAH36096	Rat PGC-3 cdna. R	C 329	14	26.9	8402	23	ABL6738	Drosophila melanog
C 257	14	26.9	3169	25	ABT18065	Aspergillus fumiga	C 330	14	26.9	11440	23	ABL17182	Drosophila melanog
C 258	14	26.9	3203	24	AAH36091	Human PGC-3a cdna.	C 331	14	26.9	11687	23	ABL27700	Drosophila melanog
C 259	14	26.9	3205	24	ABOQ3489	Human cdna seq ID	C 332	14	26.9	12775	23	ABL29294	Drosophila melanog
C 260	14	26.9	3244	23	ABL15190	Drosophila melanog	C 333	14	26.9	14395	23	ABL23912	Drosophila melanog
C 261	14	26.9	3267	24	AAH36095	Human PGC-3c cdna.	C 334	14	26.9	19867	23	ABL08912	Drosophila melanog
C 262	14	26.9	3269	25	ABT19879	Aspergillus fumiga	C 335	14	26.9	20245	23	ABL15504	Drosophila melanog
C 263	14	26.9	3333	23	ABL29285	Drosophila melanog	C 336	14	26.9	22875	23	ABL28694	Drosophila melanog
C 264	14	26.9	3385	24	AAS93265	DNA encoding novel	C 337	14	26.9	25801	22	AAH11177	Pseudomonas stutze
C 265	14	26.9	3464	24	AAS53986	cdna encoding huma	C 338	14	26.9	28906	22	AAH23705	Human tumour suppr
C 266	14	26.9	3554	23	ABL03468	Drosophila melanog	C 339	14	26.9	29316	25	ABZ66810	Orthoomeyacin biosy
C 267	14	26.9	3556	23	ABL21380	Nucleotide sequenc	C 340	14	26.9	44938	23	AAS59525	Protonibacterium
C 268	14	26.9	3585	22	AAH55492	Drosophila melanog	C 341	14	26.9	49380	23	AAH18138	Drosophila melanog
C 269	14	26.9	3602	23	ABL26683	Mammalian toxicolo	C 342	14	26.9	49617	22	AAF28541	Genomic fragment #
C 270	14	26.9	3689	22	AAS081852	Drosophila melanog	C 343	14	26.9	66788	23	AAS59515	Protonibacterium
C 271	14	26.9	3779	23	ABL04090	Orthoomeyacin biosy	C 344	14	26.9	109519	22	AAS08693	Human SURF2 genom
C 272	14	26.9	3801	25	ABZ66762	Human secreted pro	C 345	14	26.9	171936	24	ABH56565	Human cytochrome P
C 273	14	26.9	3871	20	AAH60801	Human cdna encodin	C 346	14	26.9	213456	25	AAH54072	Human cytochrome P
C 274	14	26.9	3871	22	AAS59207	Human polynucleoti	C 347	14	26.9	229354	24	ABQ74179	Shrimp white spot
C 275	14	26.9	3871	24	ABA90876	Human immunohae	C 348	14	26.9	305107	22	AAH62689	C glutamic codin
C 276	14	26.9	3953	22	AAH04659	Human immunohae	C 349	14	26.9	349980	24	ABO81846	Bifidobacterium lo
C 277	14	26.9	3953	23	AAH85377	Human testicular a	C 350	14	26.9	349980	24	ABO81847	Bifidobacterium lo
C 278	14	26.9	3953	23	ABL97566	Human immunohae	C 351	14	26.9	1230025	20	AAH91990	Nucleotide sequenc
C 279	14	26.9	3985	22	AAH85378	Drosophila melanog	C 352	13	25.0	13	23	ABF74012	Oligonucleotide SE
C 280	14	26.9	4291	23	ABL24556	Drosophila melanog	C 353	13	25.0	13	23	ABF74013	Oligonucleotide SE
C 281	14	26.9	4399	23	ABL28796	Human cdna seq ID	C 354	13	25.0	20	20	AAH94262	PCR primer used to
C 282	14	26.9	4415	24	ABOQ3381	DNA encoding novel	C 355	13	25.0	21	22	AAF96232	Human gene single
C 283	14	26.9	4502	23	AAS84550	Human oxygen regul	C 356	13	25.0	25	22	AAF25895	Human c-src/PDGFR
C 284	14	26.9	4503	18	AAH64783	Human cdna differe	C 357	13	25.0	43	22	AAF25888	Human c-src/PDGFR
C 285	14	26.9	4503	24	ABK84398	Human cdna differe	C 358	13	25.0	50	22	AAH31164	Human SNP oligonuc
C 286	14	26.9	4521	25	ABK62977	Human activated T	C 359	13	25.0	52	19	AAH43491	HIV-1 co-receptor
C 287	14	26.9	4624	23	ABL03466	Drosophila melanog	C 360	13	25.0	65	24	AAH43491	Mouse spliced tran
C 288	14	26.9	4810	24	AAS52850	Drosophila melanog	C 361	13	25.0	94	22	AAS23182	DNA encoding novel
C 289	14	26.9	4810	24	AAH17044	Bjerkandera adusta	C 362	13	25.0	115	16	AAT19383	Human gene bignacu
C 290	14	26.9	4857	23	ABL03544	Drosophila melanog	C 363	13	25.0	122	24	ABV88278	Human colon cancer
C 291	14	26.9	4941	23	ABL10270	Drosophila melanog	C 364	13	25.0	122	24	ABK47664	CDNA encoding colo
C 292	14	26.9	4944	23	AAS74937	DNA encoding novel	C 365	13	25.0	138	24	AAH89261	Human brain T calc
C 293	14	26.9	4945	23	ABL28756	Drosophila melanog	C 366	13	25.0	145	25	ABH67796	Streptococcus poly
C 294	14	26.9	5046	21	AAH1819	DNA encoding a C1	C 367	13	25.0	159	21	AAH87222	Corn ear-derived p
C 295	14	26.9	5046	23	ABA92721	Chrysosporium C1 s	C 368	13	25.0	172	25	AAH87222	Human secreted pro
C 296	14	26.9	5197	22	AAH75551	Human immune/haema	C 369	13	25.0	178	25	ABX31206	Human GDP-mannose
C 297	14	26.9	5197	22	AAH84045	Human immune/haema	C 370	13	25.0	189	25	ABX28707	Human GDP-mannose
C 298	14	26.9	5451	23	AAH25674	Drosophila melanog	C 371	13	25.0	208	22	ABX53347	Bovine foetal liver
C 299	14	26.9	5451	22	AAH25674	Pseudomonas sp lip	C 372	13	25.0	208	22	ABH70446	Human foetal liver
C 300	14	26.9	5471	24	ABH99407	Mouse ischaemic co	C 373	13	25.0	208	22	ABA37088	Probe #15554 for g

C 520	13	25.0	580	24	AB014801	Oligonucleotide fo	593	13	25.0	911	22	AAK13708	Human brain expres
C 521	13	25.0	580	24	AB014801	Oligonucleotide fo	594	13	25.0	911	22	AAK13949	Human bone marrow
C 522	13	25.0	589	21	AAF08946	Fusarium venenatum	595	13	25.0	911	22	AAI20259	Probe #10192 for g
C 523	13	25.0	592	22	AAKS8839	Human immune/haema	596	13	25.0	911	22	AAI145461	Probe #14147 used
C 524	13	25.0	595	22	AA008504	Human secreted pro	597	13	25.0	911	22	AAI05963	Probe #5954 used t
C 525	13	25.0	599	21	AACT7868	Human cancer assoc	598	13	25.0	911	23	ABS139040	Human liver single
C 526	13	25.0	609	24	ABK65542	Helicobacter pylor	599	13	25.0	911	24	ABS13538	Human genome-deriv
C 527	13	25.0	612	22	AA008528	Human secreted pro	600	13	25.0	919	21	AAK54383	Arabidopsis thalia
C 528	13	25.0	612	24	ABK78000	Bacillus clausii g	601	13	25.0	919	23	AAK593485	DNA encoding novel
C 529	13	25.0	621	24	AB047958	Oligonucleotide fo	602	13	25.0	924	23	AAK80363	DNA encoding novel
C 530	13	25.0	621	24	AB047959	Oligonucleotide fo	603	13	25.0	933	24	ABN98607	Arabidopsis thalia
C 531	13	25.0	623	21	AAAF09210	Fusarium venenatum	604	13	25.0	939	22	ABA08963	Human organic anto
C 532	13	25.0	638	22	AAH07612	Human CDNA clone (605	13	25.0	957	21	AAK51026	Arabidopsis thalia
C 533	13	25.0	638	24	ABK65669	Helicobacter pylor	606	13	25.0	957	22	ABZ14724	Arabidopsis thalia
C 534	13	25.0	640	22	AAH42840	Nucleotide sequenc	607	13	25.0	966	22	ABA46503	Human breast cell
C 535	13	25.0	642	23	ABL29545	Drosophila melanog	608	13	25.0	966	22	ABA64784	Human foetal liver
C 536	13	25.0	645	23	ABL24859	Drosophila melanog	609	13	25.0	966	22	ABA31902	Probe #10368 for g
C 537	13	25.0	651	21	AAFI2695	Aspergillus oryzae	610	13	25.0	966	22	AAK13216	Human brain expres
C 538	13	25.0	652	24	AB060365	Human colon cancer	611	13	25.0	966	22	AAK36947	Human bone marrow
C 539	13	25.0	659	19	AAV31299	E. coli J96 pathog	612	13	25.0	966	22	AAI19757	Probe #9680 for ge
C 540	13	25.0	669	24	AB065850	Arabidopsis thalia	613	13	25.0	966	22	AAI44953	Probe #13639 used
C 541	13	25.0	673	24	ABZ82542	Human secreted pro	614	13	25.0	966	22	AAI05474	Probe #5465 used t
C 542	13	25.0	687	21	AAK53563	Arabidopsis thalia	615	13	25.0	966	23	ABS38535	Human liver single
C 543	13	25.0	688	24	ABT09927	Human breast cance	616	13	25.0	966	24	ABS13031	Human genome-deriv
C 544	13	25.0	693	22	AAH06669	Human CDNA clone (617	13	25.0	967	22	AAH33455	Human colon cancer
C 545	13	25.0	699	24	ABK73909	Bacillus lichenifo	618	13	25.0	972	24	ABK65371	Arabidopsis thalia
C 546	13	25.0	701	21	AAFI1898	Aspergillus oryzae	619	13	25.0	974	21	AAK46464	Arabidopsis thalia
C 547	13	25.0	717	23	ABL28303	Drosophila melanog	620	13	25.0	977	21	AAK38710	Arabidopsis thalia
C 548	13	25.0	725	21	AAK37728	Arabidopsis thalia	621	13	25.0	985	21	AAZ98166	Human signal pepti
C 549	13	25.0	735	22	AAH06483	Human CDNA clone (622	13	25.0	989	21	AAA37107	Human PRO1558 (UNQ
C 550	13	25.0	735	24	ABZ12513	Arabidopsis thalia	623	13	25.0	989	22	AAK54419	DNA encoding prote
C 551	13	25.0	737	22	AAH06816	Human CDNA clone (624	13	25.0	989	23	ABK43597	DNA encoding novel
C 552	13	25.0	738	22	AAH67258	C glutamicum codin	625	13	25.0	989	24	AAK42264	Human ELG1 gene co
C 553	13	25.0	764	21	AAK93941	Human pancreatic c	626	13	25.0	989	24	ABK40264	DNA encoding huma
C 554	13	25.0	769	20	AAZ17031	Human gene express	627	13	25.0	994	23	ABK43895	Aspergillus oryzae
C 555	13	25.0	781	23	ABL11139	Drosophila melanog	628	13	25.0	1002	21	AAFI3243	Drosophila melanog
C 556	13	25.0	783	21	AACT7261	Human CDNA isolate	629	13	25.0	1014	23	ABL07655	Human PRO1558 (UNQ
C 557	13	25.0	783	25	ABZ52270	Human OREF ORF2816	630	13	25.0	1016	21	AAK37813	DNA encoding novel
C 558	13	25.0	788	22	AAK91743	Aspergillus oryzae	631	13	25.0	1020	14	AAK050991	Arabidopsis thalia
C 559	13	25.0	788	22	AAK93785	Human CDNA 5'-end	632	13	25.0	1022	20	AAK28354	Human prohibitin g
C 560	13	25.0	789	24	ABK05420	Human O-methyltran	633	13	25.0	1024	24	ABK40275	krtinlei coding se
C 561	13	25.0	792	24	ABL34966	Human CDNA isolate	634	13	25.0	1024	25	ABZ83378	cDNA encoding huma
C 562	13	25.0	795	22	AAK91717	Human CDNA 5'-end	635	13	25.0	1028	22	AAK68001	Toxicologically re
C 563	13	25.0	798	21	AAK93730	Human CDNA clone r	636	13	25.0	1035	23	AAK51422	Pseudomonas aerugi
C 564	13	25.0	798	22	AAK93730	Arabidopsis thalia	637	13	25.0	1037	24	ABA05419	Human O-methyltran
C 565	13	25.0	799	18	AACT49169	Partial DNA clone	638	13	25.0	1041	21	AAK49125	Arabidopsis thalia
C 566	13	25.0	812	21	AACT6634	Human OREF ORF2189	639	13	25.0	1046	21	AAK35608	Arabidopsis thalia
C 567	13	25.0	828	22	AAK91605	Human CDNA clone r	640	13	25.0	1047	23	ABL24601	Drosophila melanog
C 568	13	25.0	833	24	ABO89452	Human prostate exp	641	13	25.0	1065	24	ABO15278	Oligonucleotide fo
C 569	13	25.0	833	24	ABO89452	Human prostate exp	642	13	25.0	1065	24	ABQ15279	Oligonucleotide fo
C 570	13	25.0	846	24	ABZ12457	Arabidopsis thalia	643	13	25.0	1067	24	ABK63622	Rat sequence diffe
C 571	13	25.0	861	22	AAAF71943	Corynebacterium gl	644	13	25.0	1071	22	AAH14873	Pseudomonas sp exp
C 572	13	25.0	868	21	AAK41744	Arabidopsis thalia	645	13	25.0	1079	22	AAH14873	Human CDNA sequenc
C 573	13	25.0	868	23	ABV22742	Human prostate exp	646	13	25.0	1080	22	AAFI1683	Human alpha-II cal
C 574	13	25.0	868	23	ABV22742	Human prostate exp	647	13	25.0	1085	22	AAH72753	Human cervical can
C 575	13	25.0	879	22	AAAF64186	Drosophila melanog	648	13	25.0	1086	22	ABA46796	Human breast cell
C 576	13	25.0	879	23	ABL19399	Secreted protein-e	649	13	25.0	1086	22	ABA47311	Human breast cell
C 577	13	25.0	879	25	ABZ73540	Human secreted pro	650	13	25.0	1086	22	ABA64676	Human foetal liver
C 578	13	25.0	885	21	AAK67145	Arabidopsis thalia	651	13	25.0	1086	22	ABA65196	Human foetal liver
C 579	13	25.0	885	23	AAK43112	DNA encoding novel	652	13	25.0	1086	22	ABA31800	Probe #10266 for g
C 580	13	25.0	885	23	AAK43112	DNA encoding novel	653	13	25.0	1086	22	ABA32297	Probe #10763 for g
C 581	13	25.0	896	21	AAAS7876	Arabidopsis glossy	654	13	25.0	1086	22	AAK13113	Human brain expres
C 582	13	25.0	897	20	AAK25149	Wheat Type 1 gluc	655	13	25.0	1086	22	AAK16113	Human brain expres
C 583	13	25.0	901	24	ABZ11552	Human polynucleoti	656	13	25.0	1086	22	AAK38842	Human bone marrow
C 584	13	25.0	903	21	AAK30759	Human CDNA encodin	657	13	25.0	1086	22	AAI19650	Human bone marrow
C 585	13	25.0	903	21	AAK30759	DNA encoding human	658	13	25.0	1086	22	AAI20167	Probe #9583 for ge
C 586	13	25.0	903	21	AAH34076	Human colon cancer	659	13	25.0	1086	22	AAI4846	Probe #13532 used
C 587	13	25.0	903	24	ABK90236	Human CDNA encodin	660	13	25.0	1086	22	AAI45367	Probe #14053 used
C 588	13	25.0	903	25	ABZ42843	Human prostate exp	661	13	25.0	1086	22	AAI05369	Probe #5360 used t
C 589	13	25.0	905	23	ABV24847	Human prostate exp	662	13	25.0	1086	22	AAI05872	Probe #5863 used t
C 590	13	25.0	911	22	ABA47406	Human breast cell	663	13	25.0	1086	23	ABS38427	Human liver single
C 591	13	25.0	911	22	ABA65594	Human foetal liver	664	13	25.0	1086	23	ABS38427	Human liver single
C 592	13	25.0	911	22	ABA32395	Probe #10861 for g	665	13	25.0	1086	23	ABS38942	Human liver single

666	13	25.0	1086	24	ABSI2923	Human genome-deriv
667	13	25.0	1086	24	ABSI1441	Human genome-deriv
668	13	25.0	1093	19	AAVS7380	Maize female-prefe
669	13	25.0	1100	24	ABO61039	Human PRO1558 prot
670	13	25.0	1100	24	ABK72955	Bacillus lichenifo
671	13	25.0	1101	24	ABQ90219	M. capsulatus gene
672	13	25.0	1102	21	AACT8083	Human cancer assec
673	13	25.0	1102	22	AAH33276	Human colon cancer
674	13	25.0	1111	21	AACT9321	Arabidopsis thalia
675	13	25.0	1111	24	ABO42004	Oligonucleotide fo
676	13	25.0	1111	24	ABO42005	Oligonucleotide fo
677	13	25.0	1117	24	ABO90462	M. capsulatus gene
678	13	25.0	1127	24	ABN98317	Arabidopsis thalia
679	13	25.0	1131	21	AAZ53167	Neisseria meningit
680	13	25.0	1131	21	AAZ53168	Neisseria meningit
681	13	25.0	1132	22	AA187018	Human polynucleoti
682	13	25.0	1151	24	AA562642	CDNA sequence #429
683	13	25.0	1155	22	AAH84585	E. coli growth and
684	13	25.0	1170	23	AA568985	DNA encoding novel
685	13	25.0	1176	23	AA565036	DNA encoding novel
686	13	25.0	1182	18	AACT68050	H. pylori transpor
687	13	25.0	1186	21	AACT8711	Arabidopsis thalia
688	13	25.0	1186	21	AACT8735	Arabidopsis thalia
689	13	25.0	1190	19	AAV19463	C. utilis H1S3 gen
690	13	25.0	1227	24	AB213428	Arabidopsis thalia
691	13	25.0	1244	24	AACT0176	Human UDP-glucuron
692	13	25.0	1257	21	AACT5163	Arabidopsis thalia
693	13	25.0	1273	25	AAV52024	B4 beta expansion (
694	13	25.0	1280	19	AAV52024	Helicobacter poly
695	13	25.0	1280	21	AACT2218	Arabidopsis thalia
696	13	25.0	1281	19	AAK13956	H. pylori GHP0 699
697	13	25.0	1291	24	ABK69704	DNA encoding human
698	13	25.0	1292	24	ABK69703	DNA encoding human
699	13	25.0	1295	21	AACT0413	Arabidopsis thalia
700	13	25.0	1298	21	ABCT3673	Arabidopsis thalia
701	13	25.0	1300	24	ABQ54468	Human ovarian anct
702	13	25.0	1302	17	AACT9789	Mouse SH3P7 gene.
703	13	25.0	1304	21	AACT4791	Arabidopsis thalia
704	13	25.0	1311	25	ABX08416	DNA encoding alpha
705	13	25.0	1311	25	ABX08418	DNA encoding alpha
706	13	25.0	1311	25	ABX08427	DNA encoding alpha
707	13	25.0	1311	25	ABX08431	DNA encoding alpha
708	13	25.0	1311	25	ABX08432	DNA encoding alpha
709	13	25.0	1311	25	ABX08433	DNA encoding alpha
710	13	25.0	1311	25	ABX08435	DNA encoding alpha
711	13	25.0	1311	25	ABX08436	DNA encoding alpha
712	13	25.0	1311	25	ABX08438	DNA encoding alpha
713	13	25.0	1311	25	ABX08441	DNA encoding alpha
714	13	25.0	1311	25	ABX08443	DNA encoding alpha
715	13	25.0	1320	21	AACT0544	Human secreted pro
716	13	25.0	1326	22	AACT1061	P. putida KT2440-a
717	13	25.0	1328	19	AAV30803	Human wild-type pr
718	13	25.0	1328	20	AAV59842	Wild type probiot
719	13	25.0	1339	24	ABU88384	Pain regulated cDN
720	13	25.0	1346	25	AACT0289	Breast cancer asso
721	13	25.0	1359	22	AACT1078	P. putida KT2440-a
722	13	25.0	1366	23	ABV24243	Human prostate exp
723	13	25.0	1369	23	AACT1440	DNA encoding novel
724	13	25.0	1374	21	AACT4793	Arabidopsis thalia
725	13	25.0	1378	25	ABX08513	DNA encoding alpha
726	13	25.0	1386	25	ABSS5921	DNA encoding alpha
727	13	25.0	1386	25	ABX08450	DNA encoding alpha
728	13	25.0	1395	20	AAK14246	Mycobacterium spec
729	13	25.0	1395	22	AAH51982	Mycobacterium tube
730	13	25.0	1401	23	AACT4777	DNA encoding novel
731	13	25.0	1401	23	AACT9323	DNA encoding novel
732	13	25.0	1408	21	AACT0427	Arabidopsis thalia
733	13	25.0	1412	21	AACT0325	Arabidopsis thalia
734	13	25.0	1422	21	AACT9228	Human secreted pro
735	13	25.0	1448	24	ABO14796	Oligonucleotide fo
736	13	25.0	1452	22	AAH67004	C glutamicum codin
737	13	25.0	1452	22	AAH67004	C glutamicum codin
738	13	25.0	1452	24	AAO30175	Human UDP-glucuron
739	13	25.0	1458	20	AAK32293	M. grisea PTJ3 gen
740	13	25.0	1467	24	ABK87052	CDNA encoding mous
741	13	25.0	1491	16	AAO86278	Cladospirium herba
742	13	25.0	1509	20	AAK29858	B. cerealis multi-
743	13	25.0	1509	23	ABU24571	Drosophila melanog
744	13	25.0	1510	22	AAH65219	C glutamicum codin
745	13	25.0	1519	22	AA505176	Taxus cuspidata ox
746	13	25.0	1540	23	AA589238	DNA encoding novel
747	13	25.0	1544	22	AA159607	Human polynucleoti
748	13	25.0	1547	20	AAZ34280	Human PRO1017 nucl
749	13	25.0	1547	20	AAK87263	CDNA clone encodin
750	13	25.0	1547	20	AAK97958	Human secreted pro
751	13	25.0	1547	21	AACT8575	Human PRO1017 (UNQ
752	13	25.0	1547	21	AAA46929	CDNA encoding nove
753	13	25.0	1547	22	AA521452	Human CDNA sequenc
754	13	25.0	1547	25	AACT03811	CDNA encoding huma
755	13	25.0	1547	25	AACT04232	Human CDNA encodin
756	13	25.0	1547	25	ABX89349	DNA encoding novel
757	13	25.0	1547	25	ABX92652	CDNA encoding huma
758	13	25.0	1552	14	AAQ50697	Gene encoding este
759	13	25.0	1552	21	AAZ94209	Human transferase
760	13	25.0	1552	22	AA11754	DNA encoding human
761	13	25.0	1552	25	ABX11569	CDNA encoding nove
762	13	25.0	1564	24	ABU90039	Human polynucleoti
763	13	25.0	1570	24	ABZ66825	Arabidopsis thalia
764	13	25.0	1575	19	AACT94206	AEPIIIA clone 63GP
765	13	25.0	1579	21	AACT94332	Human secreted pro
766	13	25.0	1608	25	ABT18242	Aspergillus fumiga
767	13	25.0	1611	25	ABT18836	Aspergillus fumiga
768	13	25.0	1620	23	AA551502	Pseudomonas aerugi
769	13	25.0	1644	23	ABU28873	Drosophila melanog
770	13	25.0	1661	22	AA157821	Human polynucleoti
771	13	25.0	1668	23	ABU23329	Drosophila melanog
772	13	25.0	1693	23	ABV23342	Human prostate exp
773	13	25.0	1693	23	ABV29086	Human prostate exp
774	13	25.0	1696	22	AACT6269	Human ovary and
775	13	25.0	1696	24	ABU39935	Trichoderma harzia
776	13	25.0	1704	22	AACT68000	Human Ras GTP 9.57
777	13	25.0	1713	23	ABU18459	Corynebacterium g1
778	13	25.0	1721	22	ABU08038	Drosophila melanog
779	13	25.0	1721	22	AAU06699	Human ovary and
780	13	25.0	1726	22	ABU08037	Human reproductive
781	13	25.0	1726	22	AAU06698	Human reproductive
782	13	25.0	1746	23	AA573192	DNA encoding novel
783	13	25.0	1771	18	AACT91473	Mycobacterium tube
784	13	25.0	1771	18	AACT91470	Mycobacterium tube
785	13	25.0	1771	19	AAV64459	M. tuberculosis im
786	13	25.0	1771	19	AAV64451	Mycobacterium tube
787	13	25.0	1771	20	AACT19261	M. tuberculosis an
788	13	25.0	1771	20	AACT19049	CDNA encoding huma
789	13	25.0	1773	24	ABK54135	Rape protox-1 cDNA
790	13	25.0	1784	18	AAV04315	Rape protox-1 cDNA
791	13	25.0	1784	18	AACT86125	Rape protox-1 cDNA
792	13	25.0	1784	22	AACT15437	Oiled rape proto
793	13	25.0	1784	22	AACT6581	Oiled rape proto
794	13	25.0	1784	22	AACT17341	Oiled rape proto
795	13	25.0	1784	24	AACT43146	Rape protox-1 cDNA
796	13	25.0	1787	20	AACT6846	PCR Bait-66K DNA s
797	13	25.0	1792	21	AACT15784	Human prostate can
798	13	25.0	1800	21	AACT55783	S. lavenulae Mct
799	13	25.0	1802	24	ABK99136	HPD-1 encoding se
800	13	25.0	1803	21	AACT50076	Sense strand of hu
801	13	25.0	1812	20	AAK34247	Mycobacterium spec
802	13	25.0	1812	22	AACT60782	Human cancer agent
803	13	25.0	1812	22	AACT60782	Human cancer agent
804	13	25.0	1857	18	AACT7342	H. pylori membrane
805	13	25.0	1860	23	ABU26413	Drosophila melanog
806	13	25.0	1867	21	AAK65511	Porcine BAC-PIGF2-
807	13	25.0	1871	22	AAK94561	Human full-length
808	13	25.0	1878	25	ABK38591	N. gonorrhoeae nuc
809	13	25.0	1883	22	AACT58739	Human polynucleoti
810	13	25.0	1912	22	AAH90112	Human bone marrow
811	13	25.0	1926	25	AACT18561	Group IIT cDNA can
				15	AACT7690	Wheat germ ACCase

812	13	25.0	1926	16	AAT04948	885	13	25.0	2522	22	AAF81726	Human pro tease and
813	13	25.0	1959	24	ABK72729	886	13	25.0	2534	23	ABL12541	Drosophila melanog
814	13	25.0	1960	23	ABL12683	887	13	25.0	2534	23	ABL26322	Drosophila melanog
815	13	25.0	1961	23	AAS83820	888	13	25.0	2535	24	ABZ78277	A. niger kex cDNA.
816	13	25.0	1965	12	AAQ11815	889	13	25.0	2536	22	AAF58303	Human GTP-binding
817	13	25.0	1990	23	AAS78913	890	13	25.0	2540	23	ABL26320	Drosophila melanog
818	13	25.0	2000	15	AAQ74367	891	13	25.0	2543	23	ABL17449	Drosophila melanog
819	13	25.0	2001	24	ABN93165	892	13	25.0	2547	11	AA056691	Serum somatotropin
820	13	25.0	2019	24	ABQ70736	893	13	25.0	2559	25	ABT20656	Aspergillus fumiga
821	13	25.0	2039	19	AAV20962	894	13	25.0	2593	22	AAH16301	Human CDNA sequenc
822	13	25.0	2047	22	AAH17480	895	13	25.0	2597	20	AAV99923	DY11 forston dycto
823	13	25.0	2048	17	AAQ67534	896	13	25.0	2609	20	AAZ28355	Kiringle1 coding se
824	13	25.0	2048	17	AAH12457	897	13	25.0	2614	11	AA066631	DNA conferring te
825	13	25.0	2048	18	AAT95836	898	13	25.0	2616	23	AA578222	Human CDNA sequenc
826	13	25.0	2048	24	ABA92490	899	13	25.0	2617	22	AAH1610	Human CDNA sequenc
827	13	25.0	2052	23	AAV12899	900	13	25.0	2621	22	ABT20058	Aspergillus fumiga
828	13	25.0	2072	20	AAV99925	901	13	25.0	2641	25	AAH17810	Human CDNA sequenc
829	13	25.0	2074	21	AAZ29114	902	13	25.0	2653	21	AAAD00338	Wheat raffinose sy
830	13	25.0	2076	22	AAAD08430	903	13	25.0	2664	23	ABL12293	Drosophila melanog
831	13	25.0	2081	22	AAS29823	904	13	25.0	2667	25	ABX63277	Human CDNA #277 d1
832	13	25.0	2081	22	AAS29825	905	13	25.0	2668	23	AACT6434	Human ORFX ORF1989
833	13	25.0	2081	22	AAS29825	906	13	25.0	2675	21	AAAC98060	Drosophila melanog
834	13	25.0	2083	22	AAS29824	907	13	25.0	2682	22	AAH33182	Human colon cancer
835	13	25.0	2085	23	ABL17165	908	13	25.0	2682	25	ABT18493	Aspergillus fumiga
836	13	25.0	2087	24	AAAD24096	909	13	25.0	2735	23	ABL03405	Drosophila melanog
837	13	25.0	2093	22	AAAD13401	910	13	25.0	2737	23	ABL24858	Drosophila melanog
838	13	25.0	2101	21	AAAD00680	911	13	25.0	2743	23	AAH16808	Drosophila melanog
839	13	25.0	2105	22	AAH29939	912	13	25.0	2753	22	AAH14335	Human CDNA sequenc
840	13	25.0	2114	22	AAH13366	913	13	25.0	2765	23	ABL29544	Drosophila melanog
841	13	25.0	2117	19	AAV59658	914	13	25.0	2765	23	ABL29544	Human ATP-dependen
842	13	25.0	2117	24	ABST73645	915	13	25.0	2770	22	AAI65270	Drosophila melanog
843	13	25.0	2150	23	ABL12905	916	13	25.0	2781	23	ABL11138	Human bone remodel
844	13	25.0	2182	24	AAAD27262	917	13	25.0	2798	24	ABST70433	Human bone remodel
845	13	25.0	2187	24	ABL65382	918	13	25.0	2800	23	ABL11378	ATG10590 CDNA seq
846	13	25.0	2192	23	AAST74338	919	13	25.0	2814	25	ABT05637	Mycobacterium tube
847	13	25.0	2194	22	AAAD12607	920	13	25.0	2817	24	ABT05637	Lipoxygenase gene.
848	13	25.0	2208	24	ABSS1382	921	13	25.0	2830	13	AAO24234	Rice lipoxygenase
849	13	25.0	2208	24	AAI40121	922	13	25.0	2839	14	ABK34789	Human CDNA for nov
850	13	25.0	2210	25	ABZ24241	923	13	25.0	2847	24	ABZ78193	A. niger aspergill1
851	13	25.0	2210	25	ABZ24241	924	13	25.0	2860	23	ABL11401	Drosophila melanog
852	13	25.0	2215	23	ABL07935	925	13	25.0	2879	22	AAAD05137	Human secreted pro
853	13	25.0	2220	22	AAH16162	926	13	25.0	2887	22	AAH73562	Transducin 18 cDNA
854	13	25.0	2222	22	AAI60525	927	13	25.0	2973	18	AAT68136	H. pylori transmem
855	13	25.0	2233	22	AAAS11755	928	13	25.0	3019	23	ABL30046	Drosophila melanog
856	13	25.0	2233	25	ABX11570	929	13	25.0	3021	23	ABL03189	Drosophila melanog
857	13	25.0	2244	23	AAS51440	930	13	25.0	3027	23	ABL27593	Drosophila melanog
858	13	25.0	2257	23	ABL24210	931	13	25.0	3035	21	AAAC6462	Streptomyces globi
859	13	25.0	2270	21	AACT5892	932	13	25.0	3042	24	ABX08937	Human ephrin A2 en
860	13	25.0	2280	22	AAAF61033	933	13	25.0	3042	24	ABV78118	Human ephrin A2 DN
861	13	25.0	2280	22	AAAF61033	934	13	25.0	3042	24	ABL27593	Human polynucleoti
862	13	25.0	2281	23	ABL07965	935	13	25.0	3042	24	ABL19336	Drosophila melanog
863	13	25.0	2306	22	AAH15498	936	13	25.0	3042	24	ABL19336	Drosophila melanog
864	13	25.0	2324	22	AAK94137	937	13	25.0	3042	24	ABL24600	Human bone marrow
865	13	25.0	2329	20	AAV99636	938	13	25.0	3111	16	AAH90072	Rat very low densi
866	13	25.0	2329	22	AAAD03460	939	13	25.0	3111	16	AAH90072	DNA encoding novel
867	13	25.0	2332	22	AAK94700	940	13	25.0	3111	16	AAH90072	Aspergillus fumiga
868	13	25.0	2353	24	ABL90005	941	13	25.0	3111	16	AAH90072	Human Zici gene.
869	13	25.0	2373	21	AAFI8310	942	13	25.0	3111	16	AAH90072	Human immune/haema
870	13	25.0	2373	24	ABK54167	943	13	25.0	3111	16	AAH90072	Drosophila melanog
871	13	25.0	2373	24	ABL90420	944	13	25.0	3111	16	AAH90072	Gene #1716 used to
872	13	25.0	2409	24	ABZ82469	945	13	25.0	3111	16	AAH90072	Arabidopsis thalia
873	13	25.0	2437	22	AAH14531	946	13	25.0	3111	16	AAH90072	Drosophila melanog
874	13	25.0	2444	22	AAK94460	947	13	25.0	3111	16	AAH90072	Human PRC3-TFE3 co
875	13	25.0	2450	22	AAH88796	948	13	25.0	3111	16	AAH90072	DNA encoding novel
876	13	25.0	2455	23	ABL07283	949	13	25.0	3111	16	AAH90072	Human caleyntenin-
877	13	25.0	2459	19	AAV35490	950	13	25.0	3111	16	AAH90072	DNA encoding novel
878	13	25.0	2459	19	AAV35490	951	13	25.0	3111	16	AAH90072	Human full-length
879	13	25.0	2459	24	ABO75502	952	13	25.0	3111	16	AAH90072	Human full-length
880	13	25.0	2490	21	AACT7446	953	13	25.0	3111	16	AAH90072	Human full-length
881	13	25.0	2496	21	ABN79878	954	13	25.0	3111	16	AAH90072	Human full-length
882	13	25.0	2503	23	AAST9028	955	13	25.0	3111	16	AAH90072	Human full-length
883	13	25.0	2507	23	ABL08781	956	13	25.0	3111	16	AAH90072	Human full-length
884	13	25.0	2517	25	ABT19087	957	13	25.0	3111	16	AAH90072	Human full-length

c 958	13	25.0	3189	23	AA551474	Pseudomonas aerugi
c 959	13	25.0	3220	22	AA534856	CDNA encoding nove
c 960	13	25.0	3239	23	ABL13684	Drosophila melanog
c 961	13	25.0	3252	23	AA585001	DNA encoding novel
c 962	13	25.0	3260	23	ABL02197	Drosophila melanog
c 963	13	25.0	3266	23	ABL16043	Drosophila melanog
c 964	13	25.0	3276	22	AAK52285	Human polynucleoti
c 965	13	25.0	3283	24	AB199253	Mouse ischaemic co
c 966	13	25.0	3285	22	AAH54777	S. epidermidis gen
c 967	13	25.0	3300	23	ABL29540	Drosophila melanog
c 968	13	25.0	3301	21	AA251329	Human RNA-associat
c 969	13	25.0	3305	23	ABL11137	Drosophila melanog
c 970	13	25.0	3307	23	ABL19010	Drosophila melanog
c 971	13	25.0	3314	22	AAH54563	S. epidermidis gen
c 972	13	25.0	3333	24	AB190281	Human polynucleoti
c 973	13	25.0	3335	24	AB211773	Human polynucleoti
c 974	13	25.0	3351	20	AAZ31552	S. rochei strain E
c 975	13	25.0	3399	23	AB109972	Drosophila melanog
c 976	13	25.0	3417	23	ABL29536	Drosophila melanog
c 977	13	25.0	3440	23	ABV21315	Human prostate exp
c 978	13	25.0	3440	23	ABV22159	Human prostate exp
c 979	13	25.0	3440	23	ABV27134	Human prostate exp
c 980	13	25.0	3440	23	ABV27998	Human prostate exp
c 981	13	25.0	3443	23	ABL12680	Drosophila melanog
c 982	13	25.0	3446	24	ABZ11772	Human polynucleoti
c 983	13	25.0	3467	22	AAFS9301	Aspergillus niger
c 984	13	25.0	3467	23	ABA92733	Aspergillus niger
c 985	13	25.0	3470	25	ABT19514	Aspergillus fumiga
c 986	13	25.0	3477	23	ABT21729	Drosophila melanog
c 987	13	25.0	3504	22	AAK52083	Human polynucleoti
c 988	13	25.0	3518	20	AAK38287	Herpes simplex vir
c 989	13	25.0	3540	23	ABL07654	Drosophila melanog
c 990	13	25.0	3556	22	AA159883	Human polynucleoti
c 991	13	25.0	3570	24	ABZ78220	A. niger kex gene.
c 992	13	25.0	3573	23	ABU05120	Drosophila melanog
c 993	13	25.0	3579	23	ABV21868	Human prostate exp
c 994	13	25.0	3585	21	AA97456	Human RalGDS (hrai
c 995	13	25.0	3598	22	AAH54124	S. epidermidis gen
c 996	13	25.0	3601	24	ABK83884	Human CDNA differe
c 997	13	25.0	3608	25	ABT17648	Aspergillus fumiga
c 998	13	25.0	3632	22	AA005133	Human secreted pro
c 999	13	25.0	3668	23	ABL23328	Drosophila melanog
c1000	13	25.0	3682	23	ABL02641	Drosophila melanog

ALIGNMENTS

```
RESULT 1
AA252247/c
ID AA252247 standard; DNA: 3416 BP.
AC
XX
AC AA252247:
XX
DT 18-JUL-2000 (first entry)
XX
DE Maize glycine-rich protein 3 partial genomic clone zmGRP3.
XX
XX Maize: glycine-rich protein 3; GRP3, regulatory element; zmGRP3;
KM root specific gene expression; root abundant gene; monocotyledon;
KW pathogen resistance; pest; herbicide; growth rate; ds.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT 1..1780 /*tag= a
FT TATA_signal 1712..1719 /*tag= b
FT CDS 1781..2560 /*tag= c
FT /*product= "GRP3"
FT /note= "Glycine-rich protein 3"
```

```
FT misc_signal 2561..3416
FT /*tag= d
FT /*label= 3'__regulatory_element
FT polyA_signal 2692..2696
FT /*tag= e
PN WO200015662-A1.
PD 23-MAR-2000.
PF 10-SEP-1999; 99WO-EP06692.
PR 11-SEP-1998; 98EP-0117251.
PA (AGRI-) AGRIC TECHNOLOGY & GENETICS GMBH.
XX Feix G, Muller D;
PI WPI; 2000-271382/23.
DR Nucleic acid sequence is used for cloning and expressing a root
PT specific or root abundant gene in a plant .
XX
PS Claim 4; Page 55-56; 60pp; English.
XX
XX The present sequence is a partial genomic clone zmGRP3 encoding
CC maize glycine-rich protein 3 (GRP3). This sequence comprises 5' and
CC 3' regulatory elements useful for cloning and
CC expressing root specific or root abundant genes in plants, especially
CC monots which provide high expression efficiency and high
CC tissue specificity. Root preferred gene expression provides several
CC advantages to plants e.g. resistance to pathogens, pests, herbicides and
CC adverse weather conditions, modification of growth rate and alteration of
CC root tissue function. This sequence also provides a means of isolating
CC related regulatory sequences of other plant species which confer root
CC specificity to genes of interest operably linked to them.
XX
SQ Sequence 3416 BP; 755 A; 841 C; 950 G; 870 T; 0 other;
XX
Query Match 36.5%; Score 19; DB 21; Length 3416;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 30 CGTGGCTGCCGCGCTGCC 48
Db 2429 CGTGGCTGCCGCGCTGCC 2411
RESULT 2
AAD50522/c
ID AAD50522 standard; DNA: 1392 BP.
AC
XX
AC AAD50522:
XX
DT 24-MAR-2003 (first entry)
XX
DE Mycobacterium avium AST4 (mav_144) glycosyl sulfoltransferase DNA.
XX
XX Mycobacterial sulphatation pathway; immune response; infection; vaccine;
KM gene therapy; glycosyl sulfoltransferase; enzyme; AST4 protein; gene; ds.
XX
XX Mycobacterium avium.
OS
XX
FH Key Location/Qualifiers
FT 1..1392 /*tag= a
FT CDS /*product= "AST4 protein"
FT
XX WO200286067-A2.
XX 31-OCT-2002.
XX
XX 19-APR-2002; 2002WO-US12374.
```

XX 20-APR-2001; 2001US-285394P.
PR 26-OCT-2001; 2001US-345953P.
XX (REGC) UNIV CALIFORNIA.
XX Bertozi C, Williams SJ, Mougous J;
XX WPI; 2001-093116/08.
DR P-PSDB; AAE32776.
XX Novel mycobacterial sulfation pathway polypeptide useful in vitro
PT cell-free assay for identifying agent that reduces the activity of the
PT polypeptide -
XX
PS Disclosure; Page 114-116; 141pp; English.
XX
XX The present invention relates to mycobacterial sulphation pathway enzymes
CC and polynucleotides encoding such proteins. Sequences of the invention
CC are useful in vitro cell-free assay for identifying agents that reduce
CC the activity of the proteins. They are useful for increasing an immune
CC response to pathogenic mycobacterium in a host. They are also useful for
CC treating mycobacterial infection in mammals. Sulphation pathway proteins
CC are useful for reducing the virility and virulence of a mycobacterium.
CC They are used as vaccines and in gene therapy. The present sequence is
CC Mycobacterium avium AST4 (may_144) glycosyl sulfotransferase DNA.
XX
SQ Sequence 1392 BP; 215 A; 485 C; 481 G; 211 T; 0 other;
Query Match 32.7%; Score 17; DB 25; Length 1392;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29 TCGTCGTCGCCGCTC 45
DB 290 TCGTCGTCGCCGCTC 274
RESULT 3
ABL19972/C
ID ABL19972 standard; DNA; 2684 BP.
XX
AC ABL19972;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 11389.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PE:SE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 11389; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
CC sequences (AB01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2684 BP; 807 A; 534 C; 541 G; 802 T; 0 other;
Query Match 32.7%; Score 17; DB 23; Length 2684;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 TTCTCAGCCTTCAT 23
DB 2101 TTCTCAGCCTTCAT 2085
RESULT 4
ABL15450/C
ID ABL15450 standard; cDNA; 4310 BP.
XX
AC ABL15450;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 40832.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PE:SE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR P-PSDB; ABB71347.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 40832; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
CC sequences (AB01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 4310 BP; 1181 A; 938 C; 946 G; 1245 T; 0 other;

Query Match 32.7%; Score 17; DB 23; Length 4310;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TTCTCTACGCTCTCAT 23
|||||
DB 848 TTCTCTACGCTCTCAT 832

RESULT 5

AAA31959
ID AAA31959 standard; DNA; 276 BP.

XX AAA31959;

XX 05-JUL-2000 (first entry)

XX Plant microsatellite marker #920.

XX Plant microsatellite sequence; core repeat sequence; detection; probe;

KM DNA polymorphism; genome mapping; physical mapping; fingerprinting;

KM variety identification; genetic variability evaluation; primer; ss.

OS Eucalyptus grandis.

XX MO9967421-A1.

XX 29-DEC-1999.

XX 25-JUN-1999; 99WO-NZ00092.

XX 25-JUN-1998; 98US-0105307.

XX (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.

PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

PI Havukkala IJ, Bloksberg LN, Glenn M;

XX WPI; 2000-116958/10.

PT New plant microsatellite markers and associated flanking species for

PT the detection of polymorphic genetic markers -

PS Claim 1; Page 347; 392pp; English.

XX Sequences AAA31040-A32093 represent novel plant microsatellite sequences

CC and associated flanking species. The sequences comprise a central core

CC repeat sequence, especially selected from the sequences AAA32094-A32096

CC with left and right flanking sequences. The polynucleotide sequences

CC can be used in the detection of DNA polymorphisms, in genome mapping,

CC in physical mapping, in positional cloning of genes, in variety

CC identification and in evaluation of genetic variability within and

CC between plant tissues, populations, cultivars, species and species

CC groups. They may also be used to design hybridization probes for

CC oligonucleotide fingerprinting and library screening and to design

CC primers for microsatellite-primed PCR. Microsatellite markers are

CC useful to locate specific economically useful genes in plant genomes.

XX Sequence 276 BP; 60 A; 82 C; 88 G; 46 T; 0 other;

Query Match 30.8%; Score 16; DB 21; Length 276;

Best Local Similarity 100.0%; Pred. No. 27;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 CATGCTCTGCTGCT 36
|||||
DB 157 CATGCTCTGCTGCT 172

RESULT 6

AAA31675

ID AAA31675 standard; DNA; 357 BP.

XX AAA31675;

XX 05-JUL-2000 (first entry)

XX Plant microsatellite marker #636.

KM Plant microsatellite sequence; core repeat sequence; detection; probe;

KM DNA polymorphism; genome mapping; physical mapping; fingerprinting;

KM variety identification; genetic variability evaluation; primer; ss.

OS Eucalyptus grandis.

XX MO9967421-A1.

XX 29-DEC-1999.

XX 25-JUN-1999; 99WO-NZ00092.

XX 25-JUN-1998; 98US-0105307.

XX (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.

PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

PI Havukkala IJ, Bloksberg LN, Glenn M;

XX WPI; 2000-116958/10.

PT New plant microsatellite markers and associated flanking species for

PT the detection of polymorphic genetic markers -

PS Claim 1; Page 264; 392pp; English.

XX Sequences AAA31040-A32093 represent novel plant microsatellite sequences

CC and associated flanking species. The sequences comprise a central core

CC repeat sequence, especially selected from the sequences AAA32094-A32096

CC with left and right flanking sequences. The polynucleotide sequences

CC can be used in the detection of DNA polymorphisms, in genome mapping,

CC in physical mapping, in positional cloning of genes, in variety

CC identification and in evaluation of genetic variability within and

CC between plant tissues, populations, cultivars, species and species

CC groups. They may also be used to design hybridization probes for

CC oligonucleotide fingerprinting and library screening and to design

CC primers for microsatellite-primed PCR. Microsatellite markers are

CC useful to locate specific economically useful genes in plant genomes.

XX Sequence 357 BP; 78 A; 92 C; 109 G; 78 T; 0 other;

Query Match 30.8%; Score 16; DB 21; Length 357;

Best Local Similarity 100.0%; Pred. No. 27;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 CATGCTCTGCTGCT 36
|||||
DB 154 CATGCTCTGCTGCT 169

RESULT 7

AAA31705

AAA31705 standard; DNA; 438 BP.

XX AAA31705;

XX 05-JUL-2000 (first entry)

XX Plant microsatellite marker #666.

XX Plant microsatellite sequence; core repeat sequence; detection; probe;

KM DNA polymorphism; genome mapping; physical mapping; fingerprinting;

KM variety identification; genetic variability evaluation; primer; ss.

OS Eucalyptus grandis.


```
XX PN WO967421-A1.
XX
XX PD 29-DEC-1999.
XX
XX PF 25-JUN-1999; 99WO-NZ00092.
XX
XX PR 25-JUN-1999; 98US-0105307.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
PI Havukkala IJ, Bloksberg LN, Glenn M;
PI MPI; 2000-116958/10.
DR
XX
XX PT New plant microsatellite markers and associated flanking species for
XX the detection of polymorphic genetic markers -
XX
XX PS Claim 1; Page 273; 392pp; English.
XX
CC Sequences AAA31040-A32093 represent novel plant microsatellite sequences
CC and associated flanking species. The sequences comprise a central core
CC repeat sequence, especially selected from the sequences AAA32094-A32096
CC with left and right flanking sequences. The polynucleotide sequences
CC can be used in the detection of DNA polymorphisms, in genome mapping,
CC in physical mapping, in positional cloning of genes, in variety
CC identification and in evaluation of genetic variability within and
CC between plant tissues, populations, cultivars, species and species
CC groups. They may also be used to design hybridization probes for
CC oligonucleotide fingerprinting and library screening and to design
CC primers for microsatellite-primed PCR. Microsatellite markers are
CC useful to locate specific economically useful genes in plant genomes.
XX
SQ Sequence 438 BP; 100 A; 107 C; 125 G; 106 T; 0 other;
Query Match 30.8%; Score 16; DB 21; Length 438;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 CATCGCTCTCGTGGCT 36
DB 149 CATCGCTCTCGTGGCT 164
RESULT 8
ID AAS84162/C
XX AAS84162 standard; cDNA; 598 BP.
XX
AC AAS84162;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #19966.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Dymenac RT, Liu C, Tang YT;
XX
```

```
DR MPI; 2001-639362/73.
DR P-PSDB; ABG19975.
XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX PS Claim 1; SEQ ID NO 19966; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 598 BP; 144 A; 166 C; 171 G; 117 T; 0 other;
Query Match 30.8%; Score 16; DB 23; Length 598;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 CTCATCGCTCTCGTCG 34
DB 63 CTCATCGCTCTCGTCG 48
RESULT 9
ID AAC35333/C
XX AAC35333 standard; DNA; 638 BP.
XX
AC AAC35333;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9808.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
```


PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144614.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147360.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147306.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.

```
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 30.8%; Score 16; DB 21; Length 638;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAGTCTCTCAGC 16
Db 609 ATGCAGTCTCTCAGC 594

```
RESULT 10
AAS69014
ID AAS69014 standard; cDNA; 1582 BP.
XX
AC AAS69014;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4818.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
XX
DR P-PSDB; ABG04827.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity _
XX
PS Claim 1; SEQ ID NO 4818; 103pp; English.
```

```
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
```

Sequence 1582 BP; 216 A; 603 C; 483 G; 279 T; 1 other;

Query Match 30.8%; Score 16; DB 23; Length 1582;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CGCTCCGCGCCTCGCC 48
Db 1380 CGCTCCGCGCCTCGCC 1395

```
RESULT 11
AAD28124
ID AAD28124 standard; DNA; 1653 BP.
XX
AC AAD28124;
XX
DT 22-APR-2002 (first entry)
XX
DE Human leucine-rich repeat-8 (ZLR8) DNA #2.
XX
KW Human; leucine-rich repeat-8; ZLR8; cytosstatic; gene therapy; leukaemia;
KW endometrial adenocarcinoma; renal cell; colon; prostate; retinoblastoma;
KW cell growth disorder; glioblastoma; neuroblastoma; kidney tumour; lung;
KW mammary gland; germ cell; bladder; oesophagus; pancreas; animal feed;
KW rhabdomyosarcoma; genitourinary tract; chromosome 11q13; ds.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 1..1653
XX FT /*tag= a
XX FT /product= "ZLR8 protein"
XX FT /note= "CDS does not include stop codon"
XX FT /partial
XX FT sig_peptide 1..54
XX FT /*tag= b
XX FT mat_peptide 55..1653
XX FT /*tag= c
XX FT /product= "Mature ZLR8 protein"
XX
XX PN WO200202604-A2.
XX
XX PD 10-JAN-2002.
XX
XX PF 02-JUL-2001; 2001WO-US20999.
XX
XX PR 30-JUN-2000; 2000US-215446P.
XX
XX PA (ZYMO ) ZYMOGENETICS INC.
```

```
XX Thayer EC, Sheppard PO, Presnell SR;
PI WPI: 2002-154725/20.
XX P-PSDB; AAE17484.
DR
XX New leucine-rich repeat proteins and polynucleotides, useful for
PT diagnosing and treating disorders related to abnormal cell growth e.g.
PT retinoblastoma, renal cell, endometrial adenocarcinoma, leukemia,
PT kidney and lung tumors -
XX
XX Claim 19: Page 64-67; 82pp; English.
XX
CC The invention relates to an isolated polypeptide comprising leucine-rich
CC repeat proteins Zlir7, Zlir8, Zlir9 and spliced variants of Zlir7, Zlir9,
CC Zlir7, Zlir8, and Zlir9 proteins are useful in directing the secretion of
CC proteins of interest from a host cell and to monitor the secretion of
CC proteins in general from cells and tissues. The Zlir DNA and proteins
CC are useful in diagnosing and treating disorders related to abnormal cell
CC growth, including retinoblastoma, renal cell adenocarcinoma, endometrial
CC adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphocytic leukaemia,
CC kidney tumours, germ cell tumours, lung large cell carcinoma, mammary,
CC colon adenocarcinoma, genitourinary tract transitional cell tumours,
CC rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas
CC and prostate adenocarcinoma. Zlir protein is useful for identifying
CC agonists and antagonists of the polypeptide, for drug design, to screen
CC for cell metabolism affecting receptors, for analysis of cell phenotype,
CC and as animal feed supplement and cell culture components. Zlir DNA is
CC also useful in gene therapy. The present sequence is human Zlir8 DNA.
CC
XX Zlir8 gene is located on chromosome 11q13.
XX
SQ Sequence 1653 BP; 223 A; 617 C; 526 G; 287 T; 0 other;
XX
Query Match 30.8%; Score 16; DB 24; Length 1653;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 33 CGCTCGCGGCTCGCC 48
DB 1496 CGCTCGCGGCTCGCC 1511
RESULT 12
AAF76853
ID AAF76853 standard; cDNA; 1872 BP.
XX
XX AAF76853;
XX
DE 14-MAY-2001 (first entry)
XX
XX Human secreted protein cDNA #11.
XX
XX Human; secreted protein; immunomodulatory; antisclerotic;
XX immunosuppressive; anti-inflammatory; anti-HIV;
XX immunostimulant; cytostatic; cardiac; vascular; anti-angiogenic;
XX ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;
XX antialzheimer's; antiparkinsonian; antimicrobial; vulnereary; gene therapy;
XX immune disorder; hyperproliferative; cardiovascular; angiogenic;
XX neurological; infection; ss.
XX
XX Homo sapiens.
XX OS
XX WO200112776-A2.
XX PN
XX 22-FEB-2001.
XX PD
XX 15-AUG-2000; 2000WO-US22350.
XX PF
XX 16-AUG-1999; 99US-0148759.
XX PR
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX Shi Y, Young PE, Ebner R, Soppet DR, Ruben SM;
XX PI
```

```
XX WPI: 2001-244245/25.
DR P-PSDB; AAB70072.
XX
XX Nucleic acids encoding 18 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. Cancers, Parkinson's
PT disease and diabetic retinopathy -
XX
XX Claim 1; Page 354-355; 380pp; English.
XX
CC The present sequence is one of 18 nucleic acid molecules encoding novel
CC human secreted proteins. The nucleic acids and proteins may be used in
CC the prevention, diagnosis and treatment of diseases including immune
CC disorders (e.g. multiple sclerosis, systemic lupus erythematosus and
CC human immunodeficiency virus (HIV) infections), hyperproliferative
CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
CC (e.g. Scleritar syndrome, Chaga's cardiomyopathy, and coronary
CC arteriosclerosis), angiogenic disorders (e.g. corneal graft
CC neovascularisation and diabetic retinopathy), neurological disorders
CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
CC infectious diseases and/or for promoting wound healing, regeneration
CC and/or chemotaxis. The nucleic acid molecules may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples. The polypeptides may also be
CC used as antigens in the production of antibodies and in assays to
CC identify modulators of protein expression and activity.
XX
SQ Sequence 1872 BP; 290 A; 659 C; 604 G; 319 T; 0 other;
XX
Query Match 30.8%; Score 16; DB 22; Length 1872;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 33 CGCTCGCGGCTCGCC 48
DB 1153 CGCTCGCGGCTCGCC 1168
RESULT 13
AAD38695
ID AAD38695 standard; cDNA; 2369 BP.
XX
XX AAD38695;
XX
DE 23-SEP-2002 (first entry)
XX
XX Human LP220 secreted protein encoding cDNA.
XX
XX Human; secreted protein; atherosclerosis; Alzheimer's disease; LP220;
XX diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
XX rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
XX reperfusion injury; arteriosclerosis; wound healing; transgenic animal;
XX gene therapy; neoplasm; transgenic; poortasia; ischemia; carcinoma;
XX chromosome 11q13; gene; ss.
XX
XX Homo sapiens.
XX OS
XX Key Location/Qualifiers
XX FH 167..2074
XX FT /tag= a
XX FT /product= "human LP220 secreted protein"
XX FT sig_peptide 167..214
XX FT /tag= b
XX FT mat_peptide 215..2071
XX FT /tag= c
XX FT /product= "mature human LP220 secreted protein"
XX
XX WO200226801-A2.
XX PN
XX 04-APR-2002.
XX PD
XX 14-SEP-2001; 2001WO-US26026.
XX PF
XX
```

PR 28-SEP-2000; 2000US-236088P.
XX (ELIL) LILLY & CO ELI.
PA Su EW, Wang H;
P1 WPI; 2002-471259/50.
XX P-PSDB; AAE23980.
DR Novel proteins and polynucleotides of secreted proteins useful for
XX treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis,
PT diabetic retinopathy, arteriosclerosis, ischemia or reperfusion injury
PT -
XX Claim 1; Page 124-127; 145pp; English.
PS The invention relates to human secreted polypeptides designated LP095,
XX LP191, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic
CC acid molecules encoding such polypeptides. Novel secreted proteins of
CC the invention are used for treating diseases such as atherosclerosis,
CC Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis,
CC arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe
CC combined immunodeficiency, ischaemia, carcinomas, haemolytic anaemia,
CC reperfusion injury, neoplasms and cancer especially liver cancer. They
CC are also used for wound healing. Polynucleotides of the invention can
CC be used to generate transgenic animals or knock out animals, which in
CC turn, are useful in the development and screening of therapeutically
CC useful reagents for use in the treatment of diseases associated with
CC LP polypeptide associated activity. They are also used in gene therapy.
CC The present sequence is human LP220 secreted protein encoding cDNA.
CC LP220 gene is located on chromosome 11q13.
XX
SQ Sequence 2369 BP; 342 A; 855 C; 755 G; 417 T; 0 other;
Query Match 30.8%; Score 16; DB 24; Length 2369;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 33 CGCTGCGGCGCTCGCC 48
Db 1662 CGCTGCGGCGCTCGCC 1677
RESULT 14
ABV99421
ID ABV99421 standard; DNA; 2451 BP.
XX
AC ABV99421;
XX
DT 27-JAN-2003 (first entry)
XX
DE Human NOVA4b coding sequence.
XX
XX Human; anti-HIV; cytosstatic; antidiabetic; antiaschmatic; cachexia; AIDS;
KW antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;
KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
KW antifertility; cerebroprotective; gene therapy; NOVA; NOV; fertility;
KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; cardiovascular disorder;
KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.
XX
OS Homo sapiens.
XX
PN WO200272771-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US07288.
XX
PR 08-MAR-2001; 2001US-274101P.

PR 08-MAR-2001; 2001US-274194P.
PR 08-MAR-2001; 2001US-274281P.
PR 08-MAR-2001; 2001US-274322P.
PR 09-MAR-2001; 2001US-274849P.
PR 12-MAR-2001; 2001US-275235P.
PR 13-MAR-2001; 2001US-275578P.
PR 13-MAR-2001; 2001US-275579P.
PR 13-MAR-2001; 2001US-275601P.
PR 14-MAR-2001; 2001US-276000P.
PR 16-MAR-2001; 2001US-276766P.
PR 19-MAR-2001; 2001US-276994P.
PR 20-MAR-2001; 2001US-277239P.
PR 20-MAR-2001; 2001US-277321P.
PR 20-MAR-2001; 2001US-277327P.
PR 20-MAR-2001; 2001US-277338P.
PR 21-MAR-2001; 2001US-277791P.
PR 22-MAR-2001; 2001US-277833P.
PR 23-MAR-2001; 2001US-278152P.
PR 26-MAR-2001; 2001US-278694P.
PR 27-MAR-2001; 2001US-278899P.
PR 27-MAR-2001; 2001US-279036P.
PR 28-MAR-2001; 2001US-279344P.
PR 30-MAR-2001; 2001US-279959P.
PR 30-MAR-2001; 2001US-280233P.
PR 02-APR-2001; 2001US-280802P.
PR 02-APR-2001; 2001US-280822P.
PR 02-APR-2001; 2001US-280900P.
PR 04-APR-2001; 2001US-281194P.
PR 13-APR-2001; 2001US-283675P.
PR 30-APR-2001; 2001US-287424P.
PR 02-MAY-2001; 2001US-288066P.
PR 03-MAY-2001; 2001US-288342P.
PR 03-MAY-2001; 2001US-288528P.
PR 15-MAY-2001; 2001US-291190P.
PR 16-MAY-2001; 2001US-291099P.
PR 16-MAY-2001; 2001US-291240P.
PR 30-MAY-2001; 2001US-294485P.
PR 31-MAY-2001; 2001US-294889P.
PR 31-MAY-2001; 2001US-294899P.
PR 18-JUN-2001; 2001US-299027P.
PR 19-JUN-2001; 2001US-299303P.
PR 19-JUN-2001; 2001US-299310P.
PR 10-JUL-2001; 2001US-304354P.
PR 31-JUL-2001; 2001US-309198P.
PR 16-AUG-2001; 2001US-312903P.
PR 10-SEP-2001; 2001US-318462P.
PR 12-SEP-2001; 2001US-318770P.
PR 27-SEP-2001; 2001US-325430P.
PR 27-SEP-2001; 2001US-325681P.
PR 18-OCT-2001; 2001US-330380P.
PR 31-OCT-2001; 2001US-335301P.
PR 14-NOV-2001; 2001US-332172P.
PR 14-NOV-2001; 2001US-332271P.
PR 14-NOV-2001; 2001US-332372P.
PR 14-NOV-2001; 2001US-333184P.
PR 14-NOV-2001; 2001US-333372P.
PR 21-NOV-2001; 2001US-333209P.
PR 03-DEC-2001; 2001US-337426P.
PR 03-DEC-2001; 2001US-338092P.
PR 04-DEC-2001; 2001US-337185P.
PR 03-JAN-2002; 2002US-345705P.
PR 08-MAR-2002; 2002US-0093463.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Caeman SJ;
PI Boldog FL, Li L, Zernusen BD, Tchernev VT, Gangolli EA;
PI Verneet CAM, Pena CE, Burgess CE, Liu X, Szytek KA, Gorman L;
PI Spaderina SK, Voss BZ, Malyankar UM, Anderson DW, Patujarajan M;
PI Miller CE, Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY;
PI Pochart PF, Zhong M;
XX
DR WPI; 2002-732624/79.

DR P-PSDB; ABP70143.
 XX New NOVX polypeptides and polynucleotides, useful for preventing,
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 PT disorders, and asthma
 XX
 XX Claim 16: Page 262-263; 619pp: English.
 XX
 CC The present invention relates to new isolated proteins (NOVX) and their
 CC coding sequences (ABV99337-ABV99595 and ABP70049-ABP70149), where X is
 CC any number from 1 to 48. The NOVX proteins and coding sequences are
 CC useful in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, preferably a NOVX-associated disorder.
 CC The NOVX coding sequences and proteins are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
 CC disease, immune disorders, hematopoietic disorders, cardiovascular
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
 CC disturbances associated with obesity, metabolic syndrome X or wasting
 CC disorders associated with chronic diseases or various cancers. The NOVX
 CC coding sequences and proteins may also be used as targets for the
 CC identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, hematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods.
 XX
 SO Sequence 2451 BP; 363 A; 903 C; 767 G; 418 T; 0 other;
 Query Match 30.8%; Score 16; DB 24; Length 2451;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 33 CGCTGCCGCGCTCGCC 48
 Db 1966 CGCTGCCGCGCTCGCC 1961
 RESULT 15
 ABV99422
 ID ABV99422 standard; DNA; 2563 BP.
 XX
 AC ABV99422;
 XX
 DT 27-JAN-2003 (first entry)
 XX
 DE Human NOV44c coding sequence.
 XX
 XX Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
 KW antiinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
 KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
 KW antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;
 KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
 KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
 KW cell proliferation; hematopoiesis; wound healing; angiogenesis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200272771-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002MO-US07288.
 XX
 PR 08-MAR-2001; 2001US-274101P.
 PR 08-MAR-2001; 2001US-274194P.
 PR 08-MAR-2001; 2001US-274281P.
 PR 08-MAR-2001; 2001US-274322P.

PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275579P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276000P.
 PR 16-MAR-2001; 2001US-276776P.
 PR 19-MAR-2001; 2001US-276949P.
 PR 20-MAR-2001; 2001US-277239P.
 PR 20-MAR-2001; 2001US-277311P.
 PR 20-MAR-2001; 2001US-277377P.
 PR 20-MAR-2001; 2001US-277388P.
 PR 21-MAR-2001; 2001US-277919P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 26-MAR-2001; 2001US-278894P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 27-MAR-2001; 2001US-279036P.
 PR 28-MAR-2001; 2001US-279344P.
 PR 30-MAR-2001; 2001US-279959P.
 PR 30-MAR-2001; 2001US-280233P.
 PR 02-APR-2001; 2001US-280802P.
 PR 02-APR-2001; 2001US-280822P.
 PR 02-APR-2001; 2001US-280900P.
 PR 04-APR-2001; 2001US-281194P.
 PR 13-APR-2001; 2001US-283675P.
 PR 30-APR-2001; 2001US-287424P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 15-MAY-2001; 2001US-291159P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 30-MAY-2001; 2001US-291240P.
 PR 30-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 31-MAY-2001; 2001US-294899P.
 PR 18-JUN-2001; 2001US-294897P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309198P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335381P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 14-NOV-2001; 2001US-333272P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 08-MAR-2002; 2002US-0093463.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 PI Boldog FL, Li L, Zehushev BD, Tchervnev VT, Gangolli EA;
 PI Vermet CAM, Pena CE, Burgess CE, Liu X, Spytek KA, Gorman LJ;
 PI Spaderne SK, Vose FZ, Malysankar UM, Anderson DW, Patursajan M;
 PI Miller CE, Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY;
 PI Fochart PF, Zhong M;
 XX
 DR MPI; 2002-732824/79.
 DR P-PSDB; ABP70144.
 XX
 PT New NOVX polypeptides and polynucleotides, useful for preventing,

ABV99420
 ID ABV99420 standard; DNA: 3201 BP.
 XX
 AC ABV99420;
 XX
 DT 27-JAN-2003 (first entry)
 XX
 DE Human NOV44a coding sequence.
 XX
 XX Human: anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
 KM antiinflammatory; cardiatic; haemostatic; neuroprotective; anorectic;
 KM neurotropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
 KM antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
 KM metabolic disorder; diabetes; obesity; infectious disease; anorexia;
 KM neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KM immune disorder; haematopoietic disorder; cardiovascular disorder;
 KM bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
 KM metabolic syndrome X; wasting disorder; cell differentiation;
 KM Single nucleotide polymorphism; SNP; cell proliferation; haematopoiesis;
 KM wound healing; angiogenesis; gene; ds.
 XX
 OS Homo sapiens.
 XX
 XX Key
 FT Location/Qualifiers
 FT variation
 FT replace(604,G)
 FT /*tag= a
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation
 FT replace(684,C)
 FT /*tag= b
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation
 FT /*tag= C
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation
 FT replace(1189,C)
 FT /*tag= d
 FT /standard_name= "Single nucleotide polymorphism"
 XX
 XX MO200272771-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002MO-US07288.
 XX
 PR 08-MAR-2001; 2001US-274101P.
 PR 08-MAR-2001; 2001US-274194P.
 PR 08-MAR-2001; 2001US-274281P.
 PR 08-MAR-2001; 2001US-274322P.
 PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275579P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276000P.
 PR 16-MAR-2001; 2001US-276994P.
 PR 19-MAR-2001; 2001US-276994P.
 PR 20-MAR-2001; 2001US-277239P.
 PR 20-MAR-2001; 2001US-277321P.
 PR 20-MAR-2001; 2001US-277327P.
 PR 20-MAR-2001; 2001US-277338P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 23-MAR-2001; 2001US-278133P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 25-MAR-2001; 2001US-278894P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 27-MAR-2001; 2001US-279036P.
 PR 28-MAR-2001; 2001US-279344P.
 PR 30-MAR-2001; 2001US-279995P.
 PR 30-MAR-2001; 2001US-280233P.
 PR 02-APR-2001; 2001US-280803P.
 PR 02-APR-2001; 2001US-280822P.
 PR 02-APR-2001; 2001US-280900P.
 PR 04-APR-2001; 2001US-281194P.
 PR 13-APR-2001; 2001US-283675P.
 XX

PR 30-APR-2001; 2001US-287424P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288142P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 15-MAY-2001; 2001US-291190P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 16-MAY-2001; 2001US-291485P.
 PR 30-MAY-2001; 2001US-294899P.
 PR 31-MAY-2001; 2001US-294899P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309189P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335301P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 14-NOV-2001; 2001US-333272P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 08-MAR-2002; 2002US-0093463.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 PI Boldo FL, Li L, Zehrueen BD, Tchernev VT, Gangolli EA;
 PI Vermet CAM, Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L;
 PI Spaderna SK, Voss E2, Malyanar UM, Anderson DW, Paturajan M;
 PI Miller CE, Taupier RJ, Padigar M, Shenoy SC, Kekuda R, Gusev VV;
 PI Pochart PF, Zhong M;
 XX
 DR WPI: 2002-732824/79.
 DR P-PSDB: ABP70142.
 XX
 PT New NOVX polypeptides and polynucleotides, useful for preventing,
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 PT disorders, and asthma
 XX
 PS Claim 16; Page 261-262; 619pp; English.
 XX
 CC The present invention relates to new isolated proteins (NOVX) and their
 CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
 CC any number from 1 to 48. The NOVX proteins and coding sequences are
 CC useful in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, preferably a NOVX-associated disorder.
 CC The NOVX coding sequences and proteins are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
 CC disease, immune disorders, haematopoietic disorders, cardiovascular
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
 CC disturbances associated with obesity, metabolic syndrome X or wasting
 CC disorders associated with chronic diseases or various cancers. The NOVX
 CC coding sequences and proteins may also be used as targets for the
 CC identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods.

Sequence 3201 BP; 534 A; 1174 C; 986 G; 507 T; 0 other;

Query Match 30.8%; Score 16; DB 24; Length 3201;

Best Local Similarity 100.0%; Pred. NO. 26;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

33 CGCTCCGCGCTCCGCC 48

1966 CGCTCCGCGCTCCGCC 1981

RESULT 18

AAFI6203/C

ID AAFI6203 standard; cDNA; 3947 BP.

AAFI6203;

13-MAR-2001 (first entry)

Human prostate cancer antigen nucleotide sequence SEQ ID NO:638.

Human; prostate cancer; prostate cancer antigen; detection; diagnosis; neuroprotective; cytoskeletal; immunomodulatory; muscular; vulvovaginal; gastrointestinal; nephrotoxic; antiinfective; gynaecological; antibacterial; gene therapy; neural; immune; reproductive; renal; gastrointestinal; pulmonary; cardiovascular; proliferative disorder; wound; infectious disease; ss.

Homo sapiens.

WO200055174-A1.

21-SEP-2000.

08-MAR-2000; 2000MO-US05988.

12-MAR-1999; 99US-0124270.

(HUMA-) HUMAN GENOME SCI INC.

(ROSE/) ROSEN C A.

Rosen CA, Ruben SM;

WPI; 2000-587513/55.

P-PSDB; AAB57000.

Prostate cancer associated gene sequences, referred to as prostate cancer antigens, useful for treatment, prevention, and diagnosis of disorders such as prostate cancer -

Claim 1; Page 1079-1080; 2338pp; English.

AAFI556 to AAFI5505 encode the human prostate cancer associated proteins, called prostate cancer antigens, given in AAB56363 to AAB57302. The prostate cancer antigens can have neuroprotective, cytoskeletal, cardioprotective, immunomodulatory, muscular, vulvovaginal, gastrointestinal, nephrotoxic, antiinfective, gynaecological and antibacterial activities, and can be used in gene therapy. The prostate cancer antigen polynucleotides may be used for detection of prostate cancer, chromosome identification, as chromosome markers, and for numerous other diagnostic or research purposes. The prostate cancer antigens may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders, wounds, and infectious diseases. AAFI556 to AAFI514 to AAB57303 represent sequences used in the exemplification of the present invention.

Sequence 3947 BP; 1067 A; 979 C; 983 G; 910 T; 8 other;

Query Match 30.8%; Score 16; DB 21; Length 3947;

Best Local Similarity 100.0%; Pred. NO. 26;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

19 CTCATCGCTCTGTCG 34

72 CTCATCGCTCTGTCG 57

RESULT 19

AAH34839/C

ID AAH34839 standard; cDNA; 3953 BP.

AAH34839;

03-SEP-2001 (first entry)

Human colon cancer antigen encoding cDNA SEQ ID NO:1921.

Human; colon cancer; colon cancer antigen; diagnosis; detection; colorectal carcinoma; ss.

Homo sapiens.

WO200122920-A2.

05-APR-2001.

28-SEP-2000; 2000MO-US26524.

29-SEP-1999; 99US-0157137.

03-NOV-1999; 99US-0163280.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Barash SC, Birse CE, Rosen CA;

WPI; 2001-235357/24.

P-PSDB; AAG75434.

Nucleic acids encoding 4277 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers -

Claim 1; Page 3431-3432; 9803pp; English.

AAH2943 to AAH37195 and AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytoskeletal activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing P. Additionally, N may be used to produce the colon cancer-associated P, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent sequences used in the exemplification of the present invention.

N.B. Pages 666 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922.

Sequence 3953 BP; 1055 A; 972 C; 986 G; 930 T; 10 other;

Query Match 30.8%; Score 16; DB 22; Length 3953;

Best Local Similarity 100.0%; Pred. NO. 26;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

19 CTCATCGCTCTGTCG 34

72 CTCATCGCTCTGTCG 57

RESULT 20


```

ABLI5493
ID ABLI5493 standard; cDNA; 6507 BP.
XX
XX
AC ABLI5493;
XX
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 40961.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX P-PSDB; ABB71390.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 40961; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABLI6176-ABLI0511), expressed DNA
XX sequences (ABLI01840-ABLI6175) and the encoded proteins
XX (AAB57737-ABR72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 6507 BP; 1745 A; 1888 C; 1712 G; 1162 T; 0 other;
SQ
Query Match 30.8%; Score 16; DB 23; Length 6507;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 21 CATCGCTCTGCTGCT 36
Db 2990 CATCGCTCTGCTGCT 3005

```

```

RESULT 21
AAS84163/c
ID AAS84163 standard; cDNA; 7434 BP.
XX
XX AAS84163;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #19967.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
OS

```

```

XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDB; ABG19976.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX Claim 1; SEQ ID NO 19967; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 7434 BP; 1795 A; 2099 C; 1892 G; 1646 T; 2 other;
SQ
Query Match 30.8%; Score 16; DB 23; Length 7434;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 19 CTCATCGCTCTGCTG 34
Db 2228 CTCATCGCTCTGCTG 2213

```

```

RESULT 22
ABLI4186/c
ID ABLI4186 standard; cDNA; 15435 BP.
XX
XX ABLI4186;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 37040.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.

```

XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX P-PSDB; ABB70083.
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Claim 1; SEQ ID NO 37040; 21pp + Sequence Listing; English.
XX PS
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
XX CC sequences (ABU1840-ABU16175) and the encoded proteins
XX CC (ABBS7737-ABBS70272).
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 15435 BP; 4735 A; 3331 C; 2995 G; 4374 T; 0 other;
Query Match 30.8%; Score 16; DB 23; Length 15435;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 15 CGCTCTCATCGCTCTC 30
Db 2140 CGCTCTCATCGCTCTC 2125
RESULT 23
AAD52898
ID AAD52898 standard; DNA; 47999 BP.
XX AC
XX AAD52898;
XX DT 14-MAY-2003 (first entry)
XX DE Human tweety homologue 2 (TTYH2) gene.
XX KM Human; tweety homologue 2; TTYH2; therapy; cancer; tumour; cytosstatic;
XX KM diagnostic marker; gene; ds.
XX OS Homo sapiens.
XX FH
XX Key Location/Qualifiers
FT CDS 1946..45753
FT FT /*tag= a
FT FT /product= "Human TTYH2 protein"
FT FT 1936..2074
FT FT /*tag= b
FT FT /number= 1
FT FT 2075..10376
FT FT /*tag= c
FT FT 10377..10549
FT FT /*tag= d
FT FT /number= 2
FT FT 10550..16622
FT FT /*tag= e
FT intron

FT exon 16623..16734
FT FT /*tag= f
FT FT /number= 3
FT FT 16735..22323
FT FT /*tag= g
FT FT 23224..23444
FT FT /*tag= h
FT FT /number= 4
FT FT 23445..28299
FT FT /*tag= i
FT FT 28300..28395
FT FT /*tag= j
FT FT /number= 5
FT FT 28396..28902
FT FT /*tag= k
FT FT 28903..28975
FT FT /*tag= l
FT FT /number= 6
FT FT 28976..35372
FT FT /*tag= m
FT FT 35373..35442
FT FT /*tag= n
FT FT /number= 7
FT FT 35443..35705
FT FT /*tag= o
FT FT 35706..35761
FT FT /*tag= p
FT FT /number= 8
FT FT 35762..36266
FT FT /*tag= q
FT FT 36267..36359
FT FT /*tag= r
FT FT /number= 9
FT FT 36360..36591
FT FT /*tag= s
FT FT 36592..36684
FT FT /*tag= t
FT FT /number= 10
FT FT 36685..38529
FT FT /*tag= u
FT FT 38530..38672
FT FT /*tag= v
FT FT /number= 11
FT FT 38673..39376
FT FT /*tag= w
FT FT 39377..39562
FT FT /*tag= x
FT FT /number= 12
FT FT 39563..40050
FT FT /*tag= y
FT FT 40051..40129
FT FT /*tag= z
FT FT /number= 13
FT FT 40130..45672
FT FT /*tag= aa
FT FT 45673..47158
FT FT /*tag= ab
FT FT /number= 14
XX PN WO200292629-A1.
XX PD 21-NOV-2002.
XX XX
XX 14-MAY-2002; 2002WO-AU00591.
XX PR 14-MAY-2001; 2001AU-0004971.
XX PA (UNIV-) UNIV QUEENSLAND TECHNOLOGY.
XX PI Clements JA;
XX WPI; 2003-129264/12.
XX P-PSDB; AAE34613.
DR

XX New human tweety homolog 2 polypeptides and polynucleotides, useful for
 PT producing an antigen-binding molecule that is immuno-interactive with
 PT the polypeptide or as diagnostic markers for cancers -
 XX
 PS Claim 10; Page 128-156; 176pp; English.
 XX
 CC The invention relates to human tweety homologue 2 (TYH2) polypeptide and
 CC polynucleotide sequence. TYH2 is useful for producing an antigen-binding
 CC molecule that is immuno-interactive with the polypeptide. The agent is
 CC useful for manufacturing a medicament for restoring a normal level and/or
 CC functional activity of TYH2 expression in a patient, and for treating or
 CC preventing cancer or tumour. TYH2 sequences may also be used to provide
 CC both drug targets and regulators to promote or inhibit one or more
 CC activities, and to provide diagnostic markers for cancers. The present
 CC sequence is human TYH2 gene.
 XX
 SO Sequence 47999 BP; 11178 A; 12727 C; 12874 G; 11209 T; 11 other;
 Query Match 30.8%; Score 16; DB 25; Length 47999;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 15 CGCTCTCATCGCTCTC 30
 DB 1247 CGCTCTCATCGCTCTC 1262
 RESULT 24
 ABL15492
 ID ABL15492 standard; cDNA; 59967 BP.
 XX
 AC ABL15492;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 40958.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 PA
 PA (PEKE) PE CORP NY.
 PA
 PA Venter JC, Adams M, Li PWD, Myers EW;
 XX
 XX WPI; 2001-656860/75.
 DR P-PSDB; ABB11389.
 PT
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX Claim 1; SEQ ID NO 40958; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB85737-AB872072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/publist/pct_sequences.
 XX
 SO Sequence 59967 BP; 17305 A; 12463 C; 12291 G; 17908 T; 0 other;
 Query Match 30.8%; Score 16; DB 23; Length 59967;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 CATCGCTCTCGCTCTC 36
 DB 4640 CATCGCTCTCGCTCTC 4655
 RESULT 25
 ABQ81844/C
 ID ABQ81844 standard; DNA; 349980 BP.
 XX
 AC ABQ81844;
 XX
 DT 19-NOV-2002 (first entry)
 XX
 DE Bifidobacterium longum NCC2705 related nucleotide sequence SEQ ID:1100.
 XX
 KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
 KW anti-diarrheic; antibacterial; inhibitor of Salmonella; detection;
 KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
 KW rotavirus; food composition; pharmaceutical composition; gene; ds.
 XX
 OS Bifidobacterium longum.
 OS Synthetic.
 XX
 PN EP1227152-A1.
 XX
 PD 31-JUL-2002.
 XX
 PF 30-JAN-2001; 2001EP-0102050.
 XX
 PR 30-JAN-2001; 2001EP-0102050.
 XX
 PA (NEST) SOC PROD NESTLE SA.
 XX
 DR WPI; 2002-668397/72.
 XX
 XX Novel polynucleotide comprising Bifidobacterium genome sequence useful
 PT as a probe or primer for detecting and/or identifying Bifidobacterium
 PT longum in a biological sample -
 XX
 XX Disclosure; SEQ ID 1100; 80pp; English.
 PS
 XX
 CC The present invention describes a polynucleotide (I) comprising a
 CC sequence of a Bifidobacterium genome selected from the nucleotide
 CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
 CC least 90% identity or which hybridises with the sequences given in
 CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding
 CC a fusion protein, comprising a sequence selected from 1097 sequences
 CC given in ABP65258 to ABP63154 ligated in frame to a polynucleotide
 CC encoding a heterologous polypeptide. (I) has anti-diarrheic and
 CC antibacterial activities, and can be used as an inhibitor of Salmonella.
 CC (II) (which is a probe) is useful for the detection and/or identification
 CC of Bifidobacterium longum in a biological sample. A carrier containing
 CC the lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618)
 CC can be used for preventing and/or treating diarrhoea brought about by
 CC pathogenic bacteria and/or rotavirus. The carrier is a food composition
 CC selected from milk, yogurt, curd, cheese, fermented milks, milk based
 CC fermented products, ice-creams, fermented cereal based products, milk
 CC based powders, infant formula, pet food or a pharmaceutical composition
 CC selected from tablets, liquid bacterial suspensions, dried oral
 CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
 CC (I) is useful in DNA arrays or chips to carry out analysis of the
 CC expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent
 CC Bifidobacterium related nucleotide sequences given in the Sequence

CC Listing from the present invention but not mentioned further within the
CC specification.
CC N.B. The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied by the
CC European Patent Office.
XX
SQ Sequence 349980 BP; 69975 A; 105045 C; 104394 G; 70566 T; 0 other;
Query Match 30.8%; Score 16; DB 24; Length 349980;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 TCATCGCTCTCGTCGC 35
Db 253393 TCATCGCTCTCGTCGC 253378
RESULT 26
ABL74833
ID ABL74833 standard; cDNA; 214 BP.
XX
AC ABL74833;
XX
DT 14-MAY-2002 (first entry)
XX
DE Corn tassel-derived polynucleotide (cdps) SEQ ID NO:4207.
XX
XX Corn: corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPS;
XX inheritance; characteristic; growth; development; disease resistance;
XX environmental adaptability; quality; yield; molecular marker;
XX multi-gene trait; plant breeding; corn tassel; gene; ss.
XX
XX Zea mays.
XX
XX US2001051335-A1.
XX
XX 13-DEC-2001.
XX
XX 16-APR-1999; 99US-0294093.
XX
XX 21-APR-1998; 98US-082567P.
XX
XX (LALGU/) LALGUDI R V.
XX (ITOL/) ITO L Y.
XX (SHER/) SHERMAN B K.
XX
XX Lalgudi RV, Ito LY, Sherman BK;
XX
XX WPI: 2002-163647/21.
XX
XX Novel purified corn tassel-derived polynucleotide useful for
XX determining altered gene expression, to recover regulatory elements and
XX to follow inheritance of desirable characteristics through hybrid
XX breeding programs -
XX
XX Claim 1; SEQ ID 4207; 201bp; English.
XX
XX The present sequence describes a purified corn tassel-derived
XX polynucleotide sequence (cdps) comprising a nucleic acid sequence
XX selected from those given in ABL70627 to ABL76833. The cdps sequences
XX encode corn tassel-derived polypeptides (CDPS). The cdps sequences (I)
XX can be used for determining altered gene expression, to recover
XX regulatory elements and to follow inheritance of desirable
XX characteristics through hybrid breeding programs. (I) are also useful
XX in the evaluation, and alteration of desired characteristics associated
XX with growth and development, disease resistance, environmental
XX adaptability, quality and yield, and as molecular markers for studying
XX inheritance of multi-gene traits in a plant breeding program. (I) can be
XX used to produce a tassel-specific profile of gene transcription, a
XX transcript image, to clone regulatory elements for use in transformation
XX vectors, to express a polypeptide, to identify, isolate or extend
XX identical or related corn tassel nucleic acid sequences from DNA
XX libraries, in nucleic acid hybridisation or amplification technologies,

CC as query sequences to determine homology of known sequences, as probe
CC for use in Southern or Northern hybridisation, and to identify the
CC presence of and/or to determine the degree of similarity between two
CC (or more) nucleic acid sequences.
XX
SQ Sequence 214 BP; 68 A; 31 C; 50 G; 62 T; 3 other;
Query Match 28.8%; Score 15; DB 24; Length 214;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCAGTCTCTCACC 15
Db 182 ATGCAGTCTCTCACC 196
RESULT 27
ABX33704/C
ID ABX33704 standard; cDNA; 264 BP.
XX
AC ABX33704;
XX
DT 11-FEB-2003 (first entry)
XX
XX Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #15761.
XX
XX Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
XX cellular fucoseylation; glycoconjugate fucoseylation; transplant rejection;
XX arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
XX complex carbohydrate; gene replacement therapy; immunosuppressive;
XX anti-inflammatory; antiarthritic; antibacterial; cerebroprotective;
XX antiasthmatic; vasotropic.
XX
XX Homo sapiens.
XX
XX US2002110548-A1.
XX
XX 15-AUG-2002.
XX
XX 11-JUN-2001; 2001US-0878574.
XX
XX 22-NOV-1996; 96US-0753233.
XX 03-DEC-1997; 97US-0984246.
XX 09-SEP-1998; 98US-0149674.
XX 14-JUN-1999; 99US-0333177.
XX
XX (GENWY) GENETICS INST INC.
XX
XX Sullivan F, Kriz R, Kumar R;
XX
XX WPI: 2003-066673/06.
XX
XX New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
XX peptide, for manufacturing complex carbohydrates, or as targets for
XX screening GM4,6D antagonists for treating e.g. arthritis, or transplant
XX rejection -
XX
XX Disclosure; SEQ ID NO 15763; 6pp; English.
XX
XX The invention relates to a composition comprising a human GDP-mannose
XX 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
XX GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
XX in a mammalian subject and for treating or ameliorating diseases affected
XX by the level of cellular fucoseylation or diseases affected by the
XX fucoseylation of glycoconjugates. These diseases include arthritis,
XX transplant rejection, asthma, sepsis, reperfusion injury, stroke or
XX infection. The GM4,6D peptide or a polynucleotide encoding it is also
XX useful for manufacturing complex carbohydrates and as targets for
XX screening small molecule antagonists of the activity of the enzyme. The
XX polynucleotide is useful in developing an assay for defects in the
XX enzyme, as well as in gene replacement therapy. Sequences
XX ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding
XX human GM4,6D peptides of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at segdata.uspto.gov/sequence.html.

XX Sequence 264 BP; 66 A; 64 C; 71 G; 63 T; 0 other;

QY Query Match 28.8%; Score 15; DB 25; Length 264;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 CTCTCGTCGCTCGC 40
111 CTCTCGTCGCTCGC 97

RESULT 28

AB217641
ID AB217641 standard; cDNA; 335 BP.

AC AB217641;

DT 23-JAN-2003 (first entry)

XX S2 subtraction library cancer related clone SEQ ID NO:67.

XX Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW immune response; virology; immunology; microbiology; molecular biology;
KW recombinant DNA technology; gene; ss.

XX Homo sapiens.

PN WO200278516-A2.

PD 10-OCT-2002.

PF 28-MAR-2002; 2002WO-US10421.

PR 30-MAR-2001; 2001US-280255P.

PR 28-AUG-2001; 2001US-315563P.

PR 09-JAN-2002; 2002US-347313P.

PA (CORI-) CORIXA CORP.

PI Wang T, Wang S, Bangur CS, Gaiger A;

DR WPI; 2003-058387/05.

XX New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT preventing and treating cancer expressing CT or CP mRNA antigens, and
PT in virology, immunology, microbiology, molecular biology and
PT recombinant DNA techniques -

PS Claim 1; SEQ ID 67; 207bp; English.

XX ABO17575 to ABO20506 represent isolated polynucleotide (I) sequences, and
CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
CC invention. (I) and (II) have cytosstatic activity and can be used in gene
CC therapy and vaccines. (I), (II), antibodies and compositions from the
CC present invention are useful for diagnosing, preventing and treating
CC cancer, which expresses CT or CP mRNA antigens. They are useful for
CC stimulating immune response. They can also be useful in virology,
CC immunology, microbiology, molecular biology and recombinant DNA
CC techniques.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 335 BP; 77 A; 96 C; 96 G; 66 T; 0 other;

QY Query Match 28.8%; Score 15; DB 25; Length 335;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 CTCTCATCGCTCTCG 31
Db 137 CTCTCATCGCTCTCG 151

RESULT 29

ABX19200/c
ID ABX19200 standard; cDNA; 336 BP.

AC ABX19200;

DT 10-FEB-2003 (first entry)

XX Human GDP-mannose 4,6-dehydratase (GM4, 6D) DNA #1257.

XX Human; GDP-mannose 4,6-dehydratase; GM4, 6D; gene; ss; inflammation;
KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
KW complex carbohydrate; gene replacement therapy; immunosuppressive;
KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
KW antiasthmatic; vasotropic.

XX Homo sapiens.

PN US2002110548-A1.

PD 15-AUG-2002.

PF 11-JUN-2001; 2001US-0878574.

PR 22-NOV-1996; 96US-0753233.

PR 03-DEC-1997; 97US-0984246.

PR 09-SEP-1998; 98US-0149674.

PR 14-JUN-1999; 99US-0333177.

PA (GEMV) GENETICS INST INC.

PI Sullivan F, Kriz R, Kumar R;

DR WPI; 2003-066673/06.

XX New composition comprising GDP-mannose 4,6-dehydratase (GM4, 6D)
PT peptide, for manufacturing complex carbohydrates, or as targets for
PT screening GM4, 6D antagonists for treating e.g. arthritis, or transplant
PT rejection -

PS Disclosure; SEQ ID NO 1259; 6pp; English.

XX The invention relates to a composition comprising a human GDP-mannose
CC 4,6-dehydratase (GM4, 6D) peptide. The peptide is useful for identifying
CC GM4, 6D inhibitors. GM4, 6D inhibitors are useful for reducing inflammation
CC in a mammalian subject and for treating or ameliorating diseases affected
CC by the level of cellular fucosylation or diseases affected by the
CC fucosylation of glycoconjugates. These diseases include arthritis,
CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or
CC infection. The GM4, 6D peptide or a polynucleotide encoding it is also
CC useful for manufacturing complex carbohydrates and as targets for
CC screening small molecule antagonists of the activity of the enzyme. The
CC polynucleotide is useful in developing an assay for defects in the
CC enzyme, as well as in gene replacement therapy. Sequences
CC ABX17942-ABX17944 and ABX17947-ABX3716 represent DNA molecules encoding
CC human GM4, 6D peptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at segdata.uspto.gov/sequence.html.

XX Sequence 336 BP; 115 A; 49 C; 80 G; 92 T; 0 other;

QY Query Match 28.8%; Score 15; DB 25; Length 336;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 67 CTCCTCAGCTCTCAT 53

RESULT 30
AAS31315
ID AAS31315 standard; cDNA; 357 BP.

XX AAS31315;
AC
XX
DT 04-DEC-2001 (first entry)
XX
DE Human cDNA encoding a novel extracellular matrix protein, Seq ID No 129.
XX
XX
KW Human; secreted extracellular matrix protein; ss; immunomodulatory;
KW Anti-HIV; antineoplastic; antirheumatic; antisclerotic; cardiac; vascular;
KW cerebroprotective; chromolytic; antimicrobial; ophthalmic; cytoskeletal;
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;
KW Sezary syndrome; Gaucher's disease; neurological diseases;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW cardiac arrest; tachycardia; angina; infection; corneal infections;
KW wound healing; immunogen; gene therapy; antisense; food additive.

OS Homo sapiens.
XX
XX WO200155368-A1.
PN
XX
XX 02-AUG-2001.
PD
XX
PF 17-JAN-2001; 2001WO-US01348.
XX
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214866.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235483.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.


```

PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232297.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.

```

```

PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250360.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251088.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI, 2001-581633/65.
XX P-PSDB; AAU87122.
XX
XX New isolated nucleic acid encoding a protein for diagnosing,
XX preventing, treating or ameliorating medical conditions and used as
XX food additives or preservatives -
XX
XX Claim 1; SEQ ID NO 42; 837bp; English.
XX
XX The invention describes an isolated nucleic acid molecule (I) encoding a
XX novel central nervous system protein. (I) and polypeptides (II) encoded
XX by (I), are used to treat a medical conditions and in diagnosis of a
XX pathological condition. Disorders which are diagnosed or treated include
XX autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
XX disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,
XX angiogenesis, nervous system disorders e.g. Alzheimer's disease and
XX amyotrophic lateral sclerosis, infections caused by bacteria, viruses
XX e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
XX e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
XX adenocarcinomas and irritable bowel syndrome, reproductive system
XX disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
XX and pituitary dwarfism, cancers and disorders at the cellular level e.g.
XX leukaemia, disorders involving neovascularisation e.g. malignancies,
XX respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
XX acute kidney failure and blood related disorders e.g. myocardial
XX infarction. The polypeptides can also be used to aid wound healing and
XX epithelial cell proliferation, to prevent skin aging due to sunburn, to
XX maintain organs before transplantation, for supporting cell culture of
XX primary tissues, to regenerate tissues and in chemotaxis. The
XX polypeptides can also be used as a food additive or preservative to
XX increase or decrease storage capabilities, fat content, lipid, protein,
XX
XX
XX Query Match 28.8%; Score 15; DB 23; Length 357;
XX Best Local Similarity 100.0%; Pred. No. 94;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ATGCAGTTCCTCAC 15
XX |||||
XX DB 102 ATGCAGTTCCTCAC 116
XX
XX RESULT 32
XX ABO66639
XX ID ABO66639 standard; cDNA; 357 BP.
XX AC ABO66639;
XX XX
XX DT 23-AUG-2002 (first entry)
XX XX
XX DE Human polynucleotide SEQ ID NO 129.
XX XX

```


KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
 KW immunosuppressive; anti-inflammation; anti-HIV; antibacterial; cancer;
 KW antiparkinsonian; antischistosomal; antianemic; anticholesteric; cancer;
 KW antihypertensive; hepatoprotective; neuroprotective; anti-inflammatory;
 KW antileukemic; antidiabetic; anticancer; anticonvulsant; antifungal;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotoxic; gene therapy; vaccine;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2002042386-A1.
 XX
 PD 11-APR-2002.
 XX
 PF 17-JAN-2001; 2001US-0764870.
 XX
 PR 31-JAN-2000; 2000US-179065P.
 PR 04-FEB-2000; 2000US-160628P.
 PR 28-JUN-2000; 2000US-214886P.
 PR 07-JUL-2000; 2000US-216647P.
 PR 07-JUL-2000; 2000US-216880P.
 PR 11-JUL-2000; 2000US-217487P.
 PR 11-JUL-2000; 2000US-217496P.
 PR 14-JUL-2000; 2000US-218290P.
 PR 26-JUL-2000; 2000US-220963P.
 PR 14-AUG-2000; 2000US-220964P.
 PR 14-AUG-2000; 2000US-224518P.
 PR 14-AUG-2000; 2000US-224519P.
 PR 14-AUG-2000; 2000US-225267P.
 PR 14-AUG-2000; 2000US-225268P.
 PR 14-AUG-2000; 2000US-225270P.
 PR 14-AUG-2000; 2000US-225447P.
 PR 14-AUG-2000; 2000US-225757P.
 PR 14-AUG-2000; 2000US-225758P.
 PR 22-AUG-2000; 2000US-226868P.
 PR 30-AUG-2000; 2000US-228924P.
 PR 01-SEP-2000; 2000US-229287P.
 PR 01-SEP-2000; 2000US-229343P.
 PR 01-SEP-2000; 2000US-229344P.
 PR 01-SEP-2000; 2000US-229345P.
 PR 05-SEP-2000; 2000US-229509P.
 PR 05-SEP-2000; 2000US-229511P.
 PR 08-SEP-2000; 2000US-231413P.
 PR 21-SEP-2000; 2000US-234223P.
 PR 21-SEP-2000; 2000US-234274P.
 PR 25-SEP-2000; 2000US-234997P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236367P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 02-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-239935P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 XX
 XX (ROSE/) ROSEN C A.
 PA (RUBEN/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PT Rosen CA, Ruben SM, Barash SC;

XX
 DR WPI; 2002-470713/50.
 DR P-PSDB; ABP47964.
 XX
 PT New nucleic acid encoding human proteins, useful for diagnosis.
 PT treatment and prevention of e.g. osteoporosis, also related
 PT polypeptides and antibodies
 XX
 PS Claim 1; SEQ ID NO 129; 235pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (AB066521-AB066785) and proteins
 CC (ABP47846-ABP48110) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and antagonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html?docId=99990764870.
 XX
 SQ Sequence 357 BP; 104 A; 68 C; 86 G; 93 T; 6 other;
 Query Match 28.8%; Score 15; DB 24; Length 357;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 ATGCAGTTCTCTCAC 15
 Db 102 ATGCAGTTCTCTCAC 116
 RESULT 33
 AAS76408/c
 ID AAS76408 standard; cDNA; 416 BP.
 XX
 AC AAS76408;
 XX
 DT 13-FEB-2002. (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #12212.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PT Dymnac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG12221.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
PS Claim 1; SEQ ID No 12212; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA564197-AA594564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 416 BP; 74 A; 118 C; 128 G; 96 T; 0 other;
Query Match 28.8%; Score 15; DB 23; Length 416;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 CTCATCGCTCTCGTC 33
|||
Db 172 CTCATCGCTCTCGTC 158
|||
RESULT 34
AA219396/C
ID AA219396 standard; cDNA; 423 BP.
XX
XX AA219396;
AC
XX
XX 05-NOV-1999 (first entry)
DT
XX
XX M. tuberculosis antigen 5' L5ER-5 cDNA sequence.
DE
XX
XX Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW immunotherapy; diagnosis; immunisation; vaccine; infection;
KW immune response; skin test; ss.
XX
XX Mycobacterium tuberculosis.
OS
XX
XX WO9942076-A2.
PN
XX
XX 26-AUG-1999.
PD
XX
XX 17-FEB-1999; 99WO-US03268.
PF
XX
XX 05-MAY-1998; 98US-0072967.
PR
XX
XX 18-FEB-1998; 98US-0025197.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Campos-Neco A, Dillon DC, Hendrickson RC, Houghton R;
PI Lodes MJ, Reed SG, Skeiky YAM, Twardzik DR, Vedvick TS;
XX WPI; 1999-527409/44.
XX P-PSDB; AAY39201.
XX
XX New antigens from Mycobacterium tuberculosis useful in diagnostic
PT skin tests and protective or therapeutic vaccines or compositions

XX
XX Claim 11; Page 229; 299pp; English.
PS
XX
XX The present invention describes polypeptides comprising an immunogenic
CC part of a Mycobacterium tuberculosis antigen (Ag). Also described
CC are vaccines and fusion protein containing M. tuberculosis Ag's.
CC M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and
CC other polypeptides fragments, can be used in pharmaceutical compositions
CC or vaccines to generate a protective or therapeutic immune response to
CC M. tuberculosis and as reagents in skin tests for diagnosis of
CC tuberculosis. Ag can induce proliferation of, or cytokine secretion
CC by, T, B or natural killer cells and/or macrophages in
CC tuberculosis-immune subjects. AA219249 to AA219460 and AAY39083 to
CC AAY39225 are used in the exemplification of the present invention.
XX
SQ Sequence 423 BP; 95 A; 128 C; 129 G; 71 T; 0 other;
Query Match 28.8%; Score 15; DB 20; Length 423;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 30 CGTCGCTGCCGCCCT 44
|||
Db 26 CGTCGCTGCCGCCCT 12
|||
RESULT 35
AA219184/C
ID AA219184 standard; cDNA; 423 BP.
XX
XX AA219184;
AC
XX
XX 05-NOV-1999 (first entry)
DT
XX
XX M. tuberculosis recombinant antigen cDNA encoding 5' L5ER-5.
DE
XX
XX Antigen; diagnosis; detection; infection; antibody; immunisation;
KW vaccine; immunity; ss.
KW
XX
XX Mycobacterium tuberculosis.
C-
XX
XX WO9942118-A2.
PN
XX
XX 26-AUG-1999.
PD
XX
XX 17-FEB-1999; 99WO-US03265.
PF
XX
XX 05-MAY-1998; 98US-0072596.
PR
XX
XX 18-FEB-1998; 98US-0024753.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Campos-Neco A, Dillon DC, Hendrickson RC, Houghton R;
PI Lodes MJ, Reed SG, Skeiky YAM, Twardzik DR, Vedvick TS;
XX WPI; 1999-527416/44.
XX P-PSDB; AAY39058.
XX
XX New polypeptide comprising antigenic portions of M. tuberculosis
PT Claim 11a; Page 274; 323pp; English.
PS
XX
XX This invention describes novel recombinant antigens and their encoding
CC nucleic acids derived from Mycobacterium tuberculosis. The novel
CC polypeptides are useful for detecting M. tuberculosis infection in a
CC biological sample by detecting antibodies which bind with the
CC polypeptides, and are useful as vaccines for immunizing against
CC M. tuberculosis infection. The new detection methods are needed as
CC current vaccination strategies do not provide 100% immunity.
XX
SQ Sequence 423 BP; 95 A; 128 C; 129 G; 71 T; 0 other;
Query Match 28.8%; Score 15; DB 20; Length 423;

Best Local Similarity 100.0%; Pred. No. 94;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 CGTCGCTGCCGCT 44
|||||
Db 26 CGTCGCTGCCGCT 12

RESULT 36
AB062561
ID AB062561 standard; DNA; 448 BP.

XX AC AB062561;
XX
XX 16-AUG-2002 (first entry)
XX
DE Mycobacterium tuberculosis BAC vector clone RV129SP6.

XX KW Mycobacterium tuberculosis; Mycobacterium bovis; mycobacterium;
KW detection; BAC vector; bacterial artificial chromosome; tuberculosis;
XX gene; ds.

OS Mycobacterium tuberculosis.

PN WO954487-A2.

PD 28-OCT-1999.

PF 16-APR-1999; 99MO-1B00740.

PR 16-APR-1998; 98US-0060756.

XX (INSP) INST PASTEUR.

PI Cole S, Buchrieser-Brosch R, Gordon S, Billault A;

XX WPI; 2000-013262/01.

PT Isolation of polynucleotides from mycobacterial genomes, useful for
PT detection of Mycobacteria and for combating tuberculosis -

PS Claim 23; Page 44-45; 161pp; English.

XX CC The present invention describes a method for isolating a polynucleotide
XX of interest that is present or is expressed in a genome of a first
XX mycobacterium strain and that is absent or altered in a genome of a
XX second mycobacterium strain, which is different from the first strain
XX using a bacterial artificial chromosome (BAC) vector. Recombinant BAC
XX vectors, which are preferably immobilised, can be used to detect
XX mycobacterial nucleic acids (genomic DNA, cDNA or mRNA) in biological
XX samples. The polynucleotides identified are useful as probes or primers
XX for detecting a given mycobacterium of interest. By aligning the
XX polynucleotides contained in the recombinant BAC vectors it is possible
XX to physically map a polynucleotide of mycobacterial origin in a
XX biological sample. The methods and vectors from the present invention
XX are useful in providing information for combating tuberculosis. It is
XX possible to compare genomes between different strains or species and
XX their non-pathogenic strains or species counterparts. AB062492 to
XX CC AB062428 and AB81227 to AB81230 represent sequences used in the
XX exemplification of the present invention.

SO Sequence 448 BP; 67 A; 141 C; 133 G; 102 T; 5 other;

Query Match 28.8%; Score 15; DB 21; Length 448;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TTCTCTACGCTCTC 21
|||||
Db 352 TTCTCTACGCTCTC 366

RESULT 37

AA190273
ID AA190273 standard; cDNA; 459 BP.

XX AC AA190273;
XX

DT 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 10333.

XX DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation; ss.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

DR P-PSDB; AAO10342.

XX Isolated nucleic acids and polypeptides, useful for preventing
XX diagnosing and treating e.g. leukaemia, inflammation and immune
XX disorders -

PT Claim 1; SEQ ID NO 10333; 1399pp + Sequence listing; English.

XX CC The invention relates to human polynucleotides (AA19941-AA193841) and
XX the encoded proteins (AA000010-AA013910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activity/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 459 BP; 110 A; 142 C; 111 G; 95 T; 1 other;

Query Match 28.8%; Score 15; DB 22; Length 459;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 CTCTCATCGCTCTCG 31
|||||
Db 199 CTCTCATCGCTCTCG 213

RESULT 38

AA02001
ID AA02001 standard; cDNA; 471 BP.

XX AC AA02001;
XX

DT 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 1999.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX Homo sapiens.
 OS EPI033401-A2.
 PN
 XX 06-SEP-2000.
 PD
 XX 21-FEB-2000; 2000EP-0200610.
 PF
 XX 26-FEB-1999; 99US-0122487.
 PR
 XX (GEST) GENSET.
 PA
 XX Dumas Milne Edwards J, Duclert A, Giordano J;
 PI WPI; 2000-500381/45.
 DR P-PSDB; AAG01995.
 XX
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PS Claim 1; SEQ ID 1999; 71pp + CD-ROM; English.
 XX
 XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.
 XX
 XX Sequence 471 BP; 142 A; 87 C; 111 G; 131 T; 0 other;
 SQ
 XX
 XX Query Match 28.8%; Score 15; DB 21; Length 471;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGCAGTCTCTCTAC 15
 |||||
 Db 124 ATGCAGTCTCTCTAC 138
 RESULT 39
 ABA58638
 ID ABA58638 standard; DNA; 481 BP.
 XX
 AC ABA58638;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human foetal liver single exon nucleic acid probe #6943.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 PD
 XX 09-AUG-2001.
 PF
 XX 30-JAN-2001; 2001WO-US00669.
 XX

PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483447/52.
 DR
 XX
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human foetal liver -
 PS Claim 1; SEQ ID NO 6943; 639pp + sequence listing; English.
 XX
 XX The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 481 BP; 73 A; 136 C; 180 G; 92 T; 0 other;
 SQ
 XX
 XX Query Match 28.8%; Score 15; DB 22; Length 481;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 15 CGCTCTCATCGCTCT 29
 |||||
 Db 152 CGCTCTCATCGCTCT 166
 RESULT 40
 ABA27633
 ID ABA27633 standard; DNA; 481 BP.
 XX
 AC ABA27633;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Probe #6099 for gene expression analysis in human heart cell sample.
 XX
 KW Human; gene expression; heart; microarray; vascular system; probe;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 XX congenital heart disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157274-A2.
 PD
 XX 09-AUG-2001.
 PF
 XX 30-JAN-2001; 2001WO-US00666.
 PR
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI

```

XX WPI: 2001-488899/53.
DR
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 1; SEQ ID No 6099; 530bp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 481 BP; 73 A; 136 C; 180 G; 92 T; 0 other;

Query Match          28 8%; Score 15; DB 22; Length 481;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      15 CGCTCTCATCGCTCT 29
        |||||
        152 CGCTCTCATCGCTCT 166

```

Search completed: November 13, 2003, 09:52:46
 Job time : 249 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2003, 09:44:31 | Search time 52 Seconds

(without alignments)
441.383 Million cell updates/sec

Title: US-10-081-935-1

Sequence: 1 atgcagctctctcaagctctc.....cgctgcgcgcctcgccagctg 52

Scoring table:

Gapop_NUC

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	38.5	732	4	US-09-252-991A-13272
2	20	38.5	1056	4	US-09-252-991A-13107
3	16	30.8	810	4	US-09-252-991A-13647
4	16	30.8	879	4	US-09-252-991A-1602
5	16	30.8	1632	4	US-09-252-991A-3859
6	16	30.8	2673	4	US-09-252-991A-2993
7	16	30.8	2715	4	US-09-252-991A-2686
8	16	30.8	2814	4	US-09-252-991A-2875
9	15	28.8	423	4	US-09-072-596-259
10	15	28.8	438	4	US-09-252-991A-15918
11	15	28.8	448	3	US-09-060-756-60
12	15	28.8	448	4	US-09-670-314-60
13	15	28.8	762	4	US-09-252-991A-15888
14	15	28.8	885	4	US-09-252-991A-15774
15	15	28.8	1068	4	US-09-252-991A-15804
16	15	28.8	1119	4	US-09-252-991A-10768
17	15	28.8	1146	4	US-09-252-991A-10603
18	15	28.8	1197	4	US-09-620-312D-168
19	15	28.8	1260	4	US-09-252-991A-10250
20	15	28.8	1317	4	US-09-252-991A-11745
21	15	28.8	1620	4	US-09-252-991A-10368
22	15	28.8	1692	4	US-09-252-991A-11887
23	15	28.8	2553	4	US-09-252-991A-113
24	15	28.8	2664	4	US-09-252-991A-108
25	15	28.8	4198	3	US-09-586-719-11
26	15	28.8	4649	6	5183745-1
27	15	28.8	5118	4	US-08-669-785-3

Sequence 1, App1
Patent No. 5183745
Sequence 3, App1

28	C	28.8	6441	4	US-08-669-785-1	Sequence 1, App1
29	C	28.8	6443	6	5183745-5	Patent No. 5183745
30	C	28.8	31880	4	US-09-453-702B-242	Sequence 242, App
31	C	28.8	38653	4	US-09-922-445-1	Sequence 1, App1
32	C	28.8	44377	2	US-08-804-127C-7	Sequence 7, App1
33	C	28.8	44377	2	US-08-804-127C-7	Sequence 1, App1
34	C	28.8	68750	3	US-09-335-409-1	Sequence 1, App1
35	C	28.8	68750	4	US-09-568-102-1	Sequence 1, App1
36	C	28.8	68750	4	US-09-567-965-1	Sequence 1, App1
37	C	28.8	68750	4	US-09-568-480-1	Sequence 1, App1
38	C	28.8	68750	4	US-09-568-480-1	Sequence 1, App1
39	C	28.8	68750	4	US-09-568-480-1	Sequence 1, App1
40	C	28.8	68750	4	US-09-567-899-1	Sequence 1, App1
41	C	28.8	71989	4	US-09-443-501A-2	Sequence 2, App1
42	C	28.8	536165	1	US-09-214-800-1	Sequence 1, App1
43	C	28.8	4403765	3	US-09-103-840A-2	Sequence 2, App1
44	C	28.8	4403765	3	US-09-103-840A-2	Sequence 2, App1
45	C	28.8	4411529	3	US-09-103-840A-1	Sequence 1, App1
46	C	28.8	4411529	3	US-09-103-840A-1	Sequence 1, App1
47	C	26.9	234	4	US-09-397-787-36	Sequence 36, App1
48	C	26.9	258	4	US-09-252-991A-5283	Sequence 3283, App
49	C	26.9	308	4	US-09-312-283C-110	Sequence 110, App
50	C	26.9	310	3	US-09-188-930-110	Sequence 110, App
51	C	26.9	405	4	US-09-252-991A-2315	Sequence 2315, App
52	C	26.9	417	4	US-09-252-991A-12658	Sequence 12658, A
53	C	26.9	486	4	US-09-252-991A-11068	Sequence 11068, A
54	C	26.9	603	4	US-09-252-991A-1063	Sequence 1063, App
55	C	26.9	651	4	US-09-252-991A-4491	Sequence 4491, App
56	C	26.9	792	4	US-09-252-991A-7005	Sequence 7005, App
57	C	26.9	813	4	US-09-252-991A-11103	Sequence 11103, A
58	C	26.9	846	4	US-09-252-991A-9635	Sequence 9635, App
59	C	26.9	927	4	US-09-252-991A-7558	Sequence 7558, App
60	C	26.9	939	4	US-09-252-991A-3216	Sequence 3216, App
61	C	26.9	960	3	US-08-651-136C-1	Sequence 1, App1
62	C	26.9	960	4	US-09-229-911A-1	Sequence 1, App1
63	C	26.9	1002	4	US-09-252-991A-1099	Sequence 1099, App
64	C	26.9	1008	4	US-09-252-991A-10996	Sequence 10996, A
65	C	26.9	1020	4	US-09-252-991A-5960	Sequence 5960, App
66	C	26.9	1062	4	US-09-252-991A-9532	Sequence 9532, App
67	C	26.9	1110	4	US-09-252-991A-5978	Sequence 5978, App
68	C	26.9	1119	4	US-09-252-991A-10559	Sequence 10559, A
69	C	26.9	1209	4	US-09-252-991A-5892	Sequence 5892, App
70	C	26.9	1230	4	US-09-252-991A-947	Sequence 947, App
71	C	26.9	1239	4	US-09-252-991A-6622	Sequence 6622, App
72	C	26.9	1275	4	US-09-252-991A-2399	Sequence 2399, App
73	C	26.9	1287	4	US-09-252-991A-11255	Sequence 11255, A
74	C	26.9	1371	4	US-09-252-991A-6010	Sequence 6010, App
75	C	26.9	1371	4	US-09-252-991A-9717	Sequence 9717, App
76	C	26.9	1410	2	US-08-343-101A-6	Sequence 6, App1
77	C	26.9	1410	3	US-09-183-688-6	Sequence 6, App1
78	C	26.9	1410	4	US-09-519-485-6	Sequence 6, App1
79	C	26.9	1425	4	US-09-252-991A-10301	Sequence 10301, A
80	C	26.9	1449	4	US-09-252-991A-6642	Sequence 6642, App
81	C	26.9	1521	4	US-09-252-991A-989	Sequence 989, App
82	C	26.9	1530	4	US-09-252-991A-9942	Sequence 9942, App
83	C	26.9	1548	4	US-09-252-991A-4907	Sequence 4907, App
84	C	26.9	1602	4	US-09-252-991A-4836	Sequence 4836, App
85	C	26.9	1620	4	US-09-252-991A-12984	Sequence 12984, A
86	C	26.9	1647	4	US-09-252-991A-5576	Sequence 5576, App
87	C	26.9	1653	4	US-09-252-991A-5515	Sequence 5515, App
88	C	26.9	1710	4	US-09-252-991A-6299	Sequence 6299, App
89	C	26.9	1731	4	US-09-252-991A-6575	Sequence 6575, App
90	C	26.9	1800	4	US-09-252-991A-12456	Sequence 12456, A
91	C	26.9	1857	4	US-09-322-478-24	Sequence 24, App1
92	C	26.9	1965	4	US-09-252-991A-4464	Sequence 4464, App
93	C	26.9	1986	4	US-09-252-991A-10499	Sequence 10499, App
94	C	26.9	2028	4	US-09-252-991A-7091	Sequence 7091, App
95	C	26.9	2235	2	US-08-907-166-7	Sequence 7, App1
96	C	26.9	2634	4	US-09-391-340-7	Sequence 7, App1
97	C	26.9	2634	4	US-08-770-301A-12	Sequence 12, App1
98	C	26.9	2861	3	US-09-175-581-12	Sequence 12, App1
99	C	26.9	3105	4	US-09-252-991A-4398	Sequence 4398, App
100	C	26.9	3105	4	US-09-252-991A-4398	Sequence 4398, App

C 101	14	26.9	3231	1	US-08-074-121-4	Sequence 4, Appl1	174	13	25.0	1101	4	US-09-252-991A-10597	Sequence 10597, A
C 102	14	26.9	3231	5	PCT-US94-06447-4	Sequence 4, Appl1	C 175	13	25.0	1116	4	US-09-252-991A-2102	Sequence 2102, Ap
C 103	14	26.9	3500	4	US-09-220-132-139	Sequence 139, App	C 176	13	25.0	1116	4	US-09-252-991A-9937	Sequence 9937, Ap
C 104	14	26.9	3689	4	US-09-443-184-42	Sequence 42, Appl	C 177	13	25.0	1119	4	US-09-252-991A-8800	Sequence 8800, Ap
C 105	14	26.9	3993	4	US-09-865-621A-4	Sequence 4, Appl1	C 178	13	25.0	1119	4	US-09-252-991A-11028	Sequence 11028, A
C 106	14	26.9	4503	2	US-08-770-501A-2	Sequence 2, Appl1	C 179	13	25.0	1134	4	US-09-252-991A-3134	Sequence 3134, Ap
C 107	14	26.9	4503	3	US-09-175-581-2	Sequence 3, Appl1	C 180	13	25.0	1137	4	US-09-716-865-21	Sequence 21, Appl
C 108	14	26.9	4810	4	US-09-596-824-5	Sequence 5, Appl1	C 181	13	25.0	1140	4	US-09-252-991A-3026	Sequence 3026, Ap
C 109	14	26.9	4810	4	US-09-865-621A-3	Sequence 3, Appl1	C 182	13	25.0	1167	4	US-09-252-991A-10810	Sequence 10810, A
C 110	14	26.9	4810	4	US-09-885-329-5	Sequence 5, Appl1	C 183	13	25.0	1173	4	US-09-252-991A-3422	Sequence 3422, Ap
C 111	14	26.9	5066	4	US-09-548-938A-5	Sequence 5, Appl1	C 184	13	25.0	1173	4	US-09-252-991A-13284	Sequence 13284, A
C 112	14	26.9	5163	4	US-09-865-621A-7	Sequence 7, Appl1	C 185	13	25.0	1191	4	US-09-252-991A-10727	Sequence 10727, A
C 113	14	26.9	5163	4	US-09-865-621A-1	Sequence 1, Appl1	C 186	13	25.0	1218	1	US-08-351-473B-6	Sequence 6, Appl1
C 114	14	26.9	5509	4	US-09-865-621A-1	Sequence 5, Appl1	C 187	13	25.0	1224	4	US-09-252-991A-11985	Sequence 11983, A
C 115	14	26.9	6642	1	US-08-727-034-5	Sequence 6, Appl1	C 188	13	25.0	1230	4	US-09-252-991A-11985	Sequence 11985, A
C 116	14	26.9	6843	1	US-08-727-034-6	Sequence 1, Appl1	C 189	13	25.0	1231	4	US-09-252-991A-2858	Sequence 2858, Ap
C 117	13	25.0	1230025	4	US-09-198-452A-1	Sequence 1, Appl1	C 190	13	25.0	1234	4	US-09-252-991A-415	Sequence 415, App
C 118	13	25.0	20	4	US-09-198-452A-3588	Sequence 3588, Ap	C 191	13	25.0	1278	4	US-09-252-991A-11277	Sequence 11277, A
C 119	13	25.0	52	2	US-08-770-235A-51	Sequence 51, Appl	C 192	13	25.0	1281	4	US-09-252-991A-3691	Sequence 3691, Ap
C 120	13	25.0	145	4	US-09-313-294A-5682	Sequence 5682, Ap	C 193	13	25.0	1287	4	US-09-252-991A-13691	Sequence 10262, A
C 121	13	25.0	234	4	US-09-252-991A-3289	Sequence 3289, Ap	C 194	13	25.0	1290	5	PCT-US93-04392-2	Sequence 2, Appl1
C 122	13	25.0	238	4	US-09-470-191-37	Sequence 37, Appl	C 195	13	25.0	1302	4	US-08-630-915A-17	Sequence 17, Appl1
C 123	13	25.0	251	4	US-09-016-434-541	Sequence 541, App	C 196	13	25.0	1328	4	US-09-297-911-22	Sequence 22, Appl
C 124	13	25.0	251	4	US-09-252-991A-12812	Sequence 12812, A	C 197	13	25.0	1335	4	US-09-252-991A-973	Sequence 973, App
C 125	13	25.0	258	4	US-09-313-294A-388	Sequence 388, App	C 198	13	25.0	1362	4	US-09-252-991A-823	Sequence 823, App
C 126	13	25.0	281	4	US-09-313-294A-2572	Sequence 2572, Ap	C 199	13	25.0	1395	4	US-09-252-991A-16537	Sequence 16537, A
C 127	13	25.0	312	4	US-09-252-991A-12961	Sequence 12961, A	C 200	13	25.0	1416	4	US-09-252-991A-11168	Sequence 11168, A
C 128	13	25.0	342	2	US-08-473-020A-29	Sequence 29, Appl	C 201	13	25.0	1500	4	US-09-252-991A-9248	Sequence 9248, Ap
C 129	13	25.0	343	2	US-08-473-020A-2	Sequence 2, Appl1	C 202	13	25.0	1500	4	US-09-252-991A-6770	Sequence 6770, Ap
C 130	13	25.0	401	3	US-08-990-823-75	Sequence 75, Appl	C 203	13	25.0	1520	4	US-09-786-240-24	Sequence 24, Appl
C 131	13	25.0	401	4	US-09-477-135A-75	Sequence 75, Appl	C 204	13	25.0	1552	5	PCT-US93-04392-1	Sequence 423, App
C 132	13	25.0	406	4	US-08-928-799A-1	Sequence 1, Appl1	C 205	13	25.0	1559	4	US-09-252-991A-423	Sequence 16430, A
C 133	13	25.0	414	4	US-09-134-001C-1894	Sequence 1894, Ap	C 206	13	25.0	1570	4	US-09-252-991A-16430	Sequence 11088, A
C 134	13	25.0	423	4	US-09-252-991A-3261	Sequence 3261, Ap	C 207	13	25.0	1590	4	US-09-252-991A-11088	Sequence 11088, A
C 135	13	25.0	439	3	US-09-060-756-98	Sequence 98, Appl	C 208	13	25.0	1602	4	US-09-252-991A-3776	Sequence 3776, Ap
C 136	13	25.0	439	4	US-09-670-314-98	Sequence 98, Appl	C 209	13	25.0	1602	4	US-09-252-991A-13294	Sequence 13294, A
C 137	13	25.0	450	4	US-09-252-991A-8407	Sequence 8407, Ap	C 210	13	25.0	1615	4	US-09-252-991A-381	Sequence 381, App
C 138	13	25.0	483	4	US-09-252-991A-14803	Sequence 14803, A	C 211	13	25.0	1617	4	US-09-252-991A-1917	Sequence 1917, A
C 139	13	25.0	486	4	US-09-252-991A-15526	Sequence 15526, A	C 212	13	25.0	1634	4	US-09-252-991A-792	Sequence 792, App
C 140	13	25.0	513	4	US-09-252-991A-2008	Sequence 2008, Ap	C 213	13	25.0	1653	4	US-09-252-991A-1111	Sequence 1111, Ap
C 141	13	25.0	540	3	US-09-318-739A-1	Sequence 1, Appl1	C 214	13	25.0	1656	4	US-09-252-991A-413	Sequence 413, App
C 142	13	25.0	540	3	US-09-318-739A-1	Sequence 1, Appl1	C 215	13	25.0	1665	4	US-09-252-991A-16061	Sequence 16061, A
C 143	13	25.0	615	4	US-09-252-991A-10075	Sequence 10075, A	C 216	13	25.0	1671	4	US-09-252-991A-11030	Sequence 11030, A
C 144	13	25.0	618	4	US-09-252-991A-10075	Sequence 10379, A	C 217	13	25.0	1674	4	US-09-252-991A-1645	Sequence 1645, Ap
C 145	13	25.0	642	4	US-09-252-991A-376	Sequence 376, App	C 218	13	25.0	1680	4	US-09-252-991A-11917	Sequence 11917, A
C 146	13	25.0	642	4	US-09-252-991A-6029	Sequence 6029, Ap	C 219	13	25.0	1683	4	US-09-252-991A-8337	Sequence 8337, Ap
C 147	13	25.0	659	4	US-08-976-259-113	Sequence 113, App	C 220	13	25.0	1692	4	US-09-252-991A-12639	Sequence 12639, A
C 148	13	25.0	660	4	US-09-252-991A-4246	Sequence 4246, Ap	C 221	13	25.0	1695	4	US-09-252-991A-6569	Sequence 6569, Ap
C 149	13	25.0	669	4	US-09-252-991A-9289	Sequence 9289, Ap	C 222	13	25.0	1728	4	US-09-252-991A-18164	Sequence 18164, A
C 150	13	25.0	711	4	US-09-252-991A-11047	Sequence 11047, A	C 223	13	25.0	1735	4	US-09-252-991A-3588	Sequence 3588, Ap
C 151	13	25.0	755	4	US-09-252-991A-9287	Sequence 9287, Ap	C 224	13	25.0	1754	3	US-08-818-1112-13	Sequence 13, Appl
C 152	13	25.0	780	3	US-08-990-823-70	Sequence 70, Appl	C 225	13	25.0	1771	4	US-08-818-1112-13	Sequence 13, Appl
C 153	13	25.0	783	4	US-09-477-135A-70	Sequence 70, Appl	C 226	13	25.0	1771	4	US-09-056-556-13	Sequence 13, Appl
C 154	13	25.0	788	4	US-09-252-991A-15414	Sequence 15414, A	C 227	13	25.0	1771	4	US-09-072-991A-2221	Sequence 19, Appl
C 155	13	25.0	798	4	US-09-328-352-2221	Sequence 2221, Ap	C 228	13	25.0	1784	2	US-08-808-931-19	Sequence 19, Appl
C 156	13	25.0	801	4	US-09-252-991A-15476	Sequence 15476, A	C 229	13	25.0	1784	3	US-08-808-931-19	Sequence 19, Appl
C 157	13	25.0	811	4	US-09-252-991A-532	Sequence 532, App	C 230	13	25.0	1784	3	US-09-050-603A-19	Sequence 19, Appl
C 158	13	25.0	811	4	US-09-252-991A-11109	Sequence 11109, A	C 231	13	25.0	1784	3	US-09-102-420B-19	Sequence 19, Appl
C 159	13	25.0	836	3	US-08-581-148C-8	Sequence 8, Appl1	C 232	13	25.0	1785	4	US-09-252-991A-1528	Sequence 1528, Ap
C 160	13	25.0	896	3	US-09-170-496D-249	Sequence 249, App	C 233	13	25.0	1788	4	US-09-252-991A-4632	Sequence 4632, Ap
C 161	13	25.0	903	4	US-09-170-496D-271	Sequence 271, App	C 234	13	25.0	1800	4	US-09-266-965-16	Sequence 16, Appl
C 162	13	25.0	924	4	US-09-252-991A-514	Sequence 514, App	C 235	13	25.0	1806	4	US-09-252-991A-11213	Sequence 11213, A
C 163	13	25.0	942	4	US-09-252-991A-10592	Sequence 10592, A	C 236	13	25.0	1833	4	US-09-252-991A-8371	Sequence 8371, Ap
C 164	13	25.0	946	4	US-09-252-991A-10909	Sequence 10909, A	C 237	13	25.0	1883	4	US-09-620-3120-629	Sequence 629, App
C 165	13	25.0	946	4	US-09-252-991A-10966	Sequence 10966, A	C 238	13	25.0	1886	4	US-09-252-991A-9873	Sequence 9873, Ap
C 166	13	25.0	1002	4	US-09-252-991A-1860	Sequence 1860, Ap	C 239	13	25.0	1912	4	US-09-620-3120-742	Sequence 742, App
C 167	13	25.0	1020	1	US-08-114-461-1	Sequence 1, Appl1	C 240	13	25.0	1926	4	US-08-737-109-9	Sequence 9, Appl1
C 168	13	25.0	1020	1	US-08-192-156-1	Sequence 1, Appl1	C 241	13	25.0	1935	4	US-09-252-991A-9758	Sequence 9758, Ap
C 169	13	25.0	1020	1	US-08-370-789-1	Sequence 1, Appl1	C 242	13	25.0	1935	4	US-09-252-991A-8513	Sequence 8513, Ap
C 170	13	25.0	1020	1	US-09-252-991A-1798	Sequence 1798, A	C 243	13	25.0	1941	4	US-09-252-991A-12669	Sequence 12669, A
C 171	13	25.0	1050	4	US-09-252-991A-10973	Sequence 10973, A	C 244	13	25.0	1944	4	US-09-252-991A-8364	Sequence 8364, Ap
C 172	13	25.0	1086	4	US-09-252-991A-10259	Sequence 10259, A	C 245	13	25.0	1960	4	US-09-252-991A-8607	Sequence 8607, Ap
C 173	13	25.0	1093	4	US-09-575-602-14	Sequence 14, Appl	C 246	13	25.0	2000	1	US-08-351-473B-1	Sequence 1, Appl1

C 247	13	25.0	2001	4	US-09-134-001C-2628	Sequence 2628, Ap	C 320	13	25.0	18318	2	US-08-926-922-6	Sequence 6, App11
C 248	13	25.0	2048	3	US-08-776-246-1	Sequence 1, App11	C 321	13	25.0	18318	3	US-09-253-682-6	Sequence 6, App11
C 249	13	25.0	2048	3	US-08-776-251-1	Sequence 1, App11	C 322	13	25.0	18318	3	US-09-527-657-6	Sequence 6, App11
C 250	13	25.0	2072	4	US-09-218-363-5	Sequence 5, App11	C 323	13	25.0	33000	4	US-09-215-694-18	Sequence 120, App1
C 251	13	25.0	2072	4	US-09-252-991A-8332	Sequence 8332, Ap	C 324	13	25.0	36181	4	US-08-311-731A-120	Sequence 120, App
C 252	13	25.0	2117	4	US-09-149-476-158	Sequence 158, App	C 325	13	25.0	36519	3	US-08-923-137-2	Sequence 2, App11
C 253	13	25.0	2250	4	US-09-252-991A-4181	Sequence 4181, Ap	C 326	13	25.0	37030	4	US-08-311-731A-25	Sequence 25, App1
C 254	13	25.0	2256	4	US-09-252-991A-13322	Sequence 13322, A	C 327	13	25.0	50341	1	US-08-247-901C-1	Sequence 1, App11
C 255	13	25.0	2298	4	US-09-252-991A-845	Sequence 845, App	C 328	13	25.0	50341	2	US-09-075-904-1	Sequence 1, App11
C 256	13	25.0	2301	4	US-09-328-352-4104	Sequence 4104, Ap	C 329	13	25.0	50341	2	US-09-075-904-1	Sequence 1, App11
C 257	13	25.0	2329	4	US-09-411-977-1	Sequence 809, App1	C 330	13	25.0	52297	4	US-09-426-436-1	Sequence 1, App11
C 258	13	25.0	2367	4	US-09-252-991A-809	Sequence 809, App1	C 331	13	25.0	52297	4	US-08-705-557-1	Sequence 1, App11
C 259	13	25.0	2385	4	US-09-252-991A-11998	Sequence 11998, A	C 332	13	25.0	53500	4	US-09-266-965-76	Sequence 76, App1
C 260	13	25.0	2445	4	US-09-252-991A-14949	Sequence 14949, A	C 333	13	25.0	55216	4	US-09-716-865-23	Sequence 23, App1
C 261	13	25.0	2459	4	US-09-308-345A-1	Sequence 1, App11	C 334	13	25.0	80266	3	US-09-078-294-3	Sequence 3, App11
C 262	13	25.0	2478	4	US-09-352-991A-592	Sequence 592, App	C 335	13	25.0	80595	3	US-09-078-294-3	Sequence 3, App11
C 263	13	25.0	2547	4	US-09-252-991A-11271	Sequence 11271, A	C 336	13	25.0	83450	4	US-09-811-469-3	Sequence 3, App11
C 264	13	25.0	2597	4	US-09-218-363-1	Sequence 1, App11	C 337	13	25.0	536165	4	US-09-214-808-1	Sequence 1, App11
C 265	13	25.0	2798	4	US-09-484-970B-90	Sequence 90, App1	C 338	12	23.1	1664976	4	US-08-916-421B-1	Sequence 1, App11
C 266	13	25.0	2823	4	US-09-252-991A-2840	Sequence 2840, Ap	C 339	12	23.1	24	1	US-08-450-186-9	Sequence 1, App11
C 267	13	25.0	2907	4	US-09-252-991A-831	Sequence 831, App	C 340	12	23.1	24	1	US-08-450-186-9	Sequence 1, App11
C 268	13	25.0	2964	4	US-09-252-991A-10912	Sequence 10912, A	C 341	12	23.1	24	1	US-08-450-186-60	Sequence 60, App1
C 269	13	25.0	2967	4	US-09-252-991A-13128	Sequence 13128, A	C 342	12	23.1	24	1	US-08-323-257-1	Sequence 1, App11
C 270	13	25.0	3027	4	US-09-252-991A-2758	Sequence 2758, Ap	C 343	12	23.1	24	1	US-08-323-257-9	Sequence 9, App11
C 271	13	25.0	3102	4	US-09-252-991A-4429	Sequence 4429, Ap	C 344	12	23.1	24	1	US-08-323-257-9	Sequence 9, App11
C 272	13	25.0	3116	1	US-08-149-103-2	Sequence 2, App11	C 345	12	23.1	24	1	US-08-323-257-60	Sequence 60, App1
C 273	13	25.0	3116	1	US-08-451-883-2	Sequence 2, App11	C 346	12	23.1	24	1	US-08-779-341-10	Sequence 10, App1
C 274	13	25.0	3138	3	US-09-234-332-5	Sequence 5, App11	C 347	12	23.1	24	2	US-08-779-341-10	Sequence 10, App1
C 275	13	25.0	3144	4	US-09-252-991A-14688	Sequence 14688, A	C 348	12	23.1	24	3	US-08-841-267-6	Sequence 6, App11
C 276	13	25.0	3198	4	US-09-252-991A-15066	Sequence 15066, A	C 349	12	23.1	24	3	US-09-502-966-15	Sequence 15, App1
C 277	13	25.0	3204	4	US-09-252-991A-3126	Sequence 3126, Ap	C 350	12	23.1	24	3	US-09-502-966-15	Sequence 15, App1
C 278	13	25.0	3651	4	US-07-852-132A-13	Sequence 1326, Ap	C 351	12	23.1	24	5	PCT-US95-02865-10	Sequence 10, App1
C 279	13	25.0	3666	6	5248670-3	Patent No. 5248670	C 352	12	23.1	26	2	US-08-859-998-678	Sequence 10, App1
C 280	13	25.0	3823	4	US-09-453-702B-132	Sequence 132, App1	C 353	12	23.1	26	2	US-09-225-928-678	Sequence 678, App
C 281	13	25.0	3962	1	US-08-336-343A-1	Sequence 1, App11	C 354	12	23.1	26	4	US-09-825-201B-678	Sequence 678, App
C 282	13	25.0	3989	1	US-08-327-494A-1	Sequence 1, App11	C 355	12	23.1	32	1	US-08-623-891-59	Sequence 59, App1
C 283	13	25.0	3989	1	US-08-327-494A-3	Sequence 3, App11	C 356	12	23.1	32	1	US-09-340-881-59	Sequence 59, App1
C 284	13	25.0	3989	5	PCT-US95-13659-1	Sequence 1, App11	C 357	12	23.1	32	4	US-09-634-262-59	Sequence 59, App1
C 285	13	25.0	3989	5	PCT-US95-13659-3	Sequence 3, App11	C 358	12	23.1	40	3	US-08-841-267-8	Sequence 8, App11
C 286	13	25.0	4074	4	US-09-252-991A-4737	Sequence 4737, Ap	C 359	12	23.1	40	4	US-08-585-593A-16	Sequence 16, App1
C 287	13	25.0	4092	4	US-09-252-991A-4771	Sequence 4771, Ap	C 360	12	23.1	111	4	US-09-065-914B-11	Sequence 11, App1
C 288	13	25.0	4348	2	US-08-915-868-1	Sequence 1, App11	C 361	12	23.1	137	5	PCT-US95-10813-23	Sequence 23, App1
C 289	13	25.0	4377	3	US-08-911-853-28	Sequence 28, App1	C 362	12	23.1	137	5	US-08-299-498A-23	Sequence 23, App1
C 290	13	25.0	4377	3	US-09-479-409-28	Sequence 28, App1	C 363	12	23.1	175	3	US-08-990-823-14	Sequence 14, App1
C 291	13	25.0	4377	3	US-09-479-453-28	Sequence 28, App1	C 364	12	23.1	175	4	US-09-477-135A-14	Sequence 14, App1
C 292	13	25.0	4394	1	US-08-095-734-1	Sequence 1, App11	C 365	12	23.1	202	4	US-08-585-593A-19	Sequence 19, App1
C 293	13	25.0	4394	2	US-08-444-623-1	Sequence 1, App11	C 366	12	23.1	202	4	US-08-585-593A-19	Sequence 19, App1
C 294	13	25.0	4394	3	US-08-471-869-1	Sequence 1, App11	C 367	12	23.1	203	1	US-08-382-933-1	Sequence 1, App11
C 295	13	25.0	4394	4	US-09-342-563-1	Sequence 1, App11	C 368	12	23.1	203	4	US-08-585-593A-1	Sequence 1, App11
C 296	13	25.0	4394	5	PCT-US94-08267-1	Sequence 1, App11	C 369	12	23.1	213	3	US-09-060-756-259	Sequence 259, App
C 297	13	25.0	4563	4	US-09-252-991A-930	Sequence 930, App	C 370	12	23.1	213	4	US-09-670-314-259	Sequence 259, App
C 298	13	25.0	4663	4	US-09-221-017B-817	Sequence 817, App	C 371	12	23.1	241	4	US-09-389-681-370	Sequence 370, App
C 299	13	25.0	5461	3	US-07-598-289B-7	Sequence 7, App11	C 372	12	23.1	241	4	US-09-620-405B-370	Sequence 370, App
C 300	13	25.0	6114	4	US-09-620-312D-98	Sequence 98, App1	C 373	12	23.1	242	4	US-09-433-826B-370	Sequence 370, App
C 301	13	25.0	6513	1	US-08-338-702-7	Sequence 7, App11	C 374	12	23.1	242	4	US-08-981-803-12	Sequence 12, App1
C 302	13	25.0	6513	1	US-08-337-339-7	Sequence 7, App11	C 375	12	23.1	242	3	US-08-981-803-12	Sequence 12, App1
C 303	13	25.0	6513	1	US-08-337-339-7	Sequence 7, App11	C 376	12	23.1	242	3	US-08-983-440-12	Sequence 12, App1
C 304	13	25.0	6513	5	PCT-US95-14263-7	Sequence 7, App11	C 377	12	23.1	242	3	US-08-983-440-12	Sequence 12, App1
C 305	13	25.0	6513	5	PCT-US95-14378-7	Sequence 7, App11	C 378	12	23.1	242	4	US-09-367-895-12	Sequence 12, App1
C 306	13	25.0	6513	5	US-08-808-793-24	Sequence 24, App1	C 379	12	23.1	242	4	US-09-367-895-26	Sequence 26, App1
C 307	13	25.0	6611	1	US-08-402-282-2	Sequence 2, App11	C 380	12	23.1	245	4	US-09-313-294A-321	Sequence 321, App
C 308	13	25.0	6611	1	US-08-508-004-2	Sequence 2, App11	C 381	12	23.1	246	2	US-08-673-190A-9	Sequence 9, App11
C 309	13	25.0	6611	1	US-08-402-066-2	Sequence 2, App11	C 382	12	23.1	247	2	US-08-403-852D-12	Sequence 12, App1
C 310	13	25.0	6611	1	US-08-402-066-2	Sequence 2, App11	C 383	12	23.1	247	3	US-08-510-646B-12	Sequence 12, App1
C 311	13	25.0	6816	4	US-09-404-650-1	Sequence 1, App11	C 384	12	23.1	255	4	US-09-231-81B-12	Sequence 12, App1
C 312	13	25.0	6855	4	US-09-404-650-3	Sequence 1, App11	C 385	12	23.1	257	4	US-09-252-991A-16186	Sequence 16186, A
C 313	13	25.0	7176	4	US-09-421-017B-626	Sequence 626, App	C 386	12	23.1	263	3	US-09-313-294A-321	Sequence 321, App
C 314	13	25.0	7808	4	US-09-453-702B-247	Sequence 247, App	C 387	12	23.1	263	3	US-08-990-823-38	Sequence 38, App1
C 315	13	25.0	8371	4	US-09-065-027-1	Sequence 1, App11	C 388	12	23.1	264	4	US-09-477-135A-80	Sequence 80, App1
C 316	13	25.0	11561	2	US-08-450-332-1	Sequence 1, App11	C 389	12	23.1	270	4	US-09-252-991A-14104	Sequence 14104, A
C 317	13	25.0	11561	2	US-08-637-640-1	Sequence 1, App11	C 390	12	23.1	273	4	US-09-252-991A-9403	Sequence 9403, Ap
C 318	13	25.0	11561	3	US-09-004-406C-1	Sequence 1, App11	C 391	12	23.1	273	4	US-09-252-991A-9403	Sequence 9403, Ap
C 319	13	25.0	18318	1	US-08-414-926A-6	Sequence 6, App11	C 392	12	23.1	279	4	US-09-313-294A-5268	Sequence 5268, Ap

393	12	23.1	280	4	US-09-313-294A-3395	Sequence 3395, Ap	466	12	23.1	474	4	US-09-670-314-86	Sequence 86, Appl1
394	12	23.1	281	4	US-09-313-294A-3379	Sequence 3379, Ap	467	12	23.1	477	4	US-09-252-991A-14502	Sequence 14502, A
395	12	23.1	281	4	US-09-313-294A-4409	Sequence 4409, Ap	468	12	23.1	482	3	US-09-060-756-113	Sequence 113, Ap
396	12	23.1	282	4	US-09-313-294A-6360	Sequence 6360, Ap	469	12	23.1	482	4	US-09-670-314-113	Sequence 113, Ap
397	12	23.1	286	4	US-09-313-294A-987	Sequence 987, Ap	470	12	23.1	483	4	US-09-252-991A-1470	Sequence 1470, Ap
398	12	23.1	288	4	US-09-313-294A-4782	Sequence 4782, Ap	471	12	23.1	483	4	US-09-252-991A-3887	Sequence 3887, Ap
399	12	23.1	293	4	US-09-313-294A-1392	Sequence 1392, Ap	472	12	23.1	486	4	US-09-252-991A-15115	Sequence 15115, A
400	12	23.1	294	4	US-08-104-072B-3	Sequence 3, Appl1	473	12	23.1	489	3	US-09-060-756-63	Sequence 63, Appl1
401	12	23.1	300	4	US-09-313-294A-4514	Sequence 4514, Ap	474	12	23.1	489	4	US-09-670-314-63	Sequence 63, Appl1
402	12	23.1	305	4	US-09-313-294A-7029	Sequence 7029, Ap	475	12	23.1	489	4	US-09-252-991A-16020	Sequence 16020, A
403	12	23.1	312	4	US-09-252-991A-7029	Sequence 1073, Ap	476	12	23.1	489	4	US-09-252-991A-6357	Sequence 6357, Ap
404	12	23.1	316	4	US-09-165-868-6	Sequence 6, Appl1	477	12	23.1	485	4	US-09-252-991A-4430	Sequence 4430, Ap
405	12	23.1	330	4	US-09-252-991A-13785	Sequence 13785, A	478	12	23.1	488	4	US-09-455-960-3	Sequence 3, Appl1
406	12	23.1	336	4	US-09-252-991A-13036	Sequence 13036, A	479	12	23.1	488	4	US-09-252-991A-1030	Sequence 1030, Ap
407	12	23.1	332	4	US-09-252-991A-12505	Sequence 12505, A	480	12	23.1	505	4	US-09-470-191-56	Sequence 56, Appl1
408	12	23.1	357	4	US-09-252-991A-441	Sequence 441, Ap	481	12	23.1	510	4	US-09-252-991A-11546	Sequence 11546, A
409	12	23.1	360	4	US-09-252-991A-5974	Sequence 5974, Ap	482	12	23.1	513	3	US-09-385-982-302	Sequence 302, Ap
410	12	23.1	360	4	US-09-252-991A-8242	Sequence 8242, Ap	483	12	23.1	513	3	US-09-252-991A-5290	Sequence 5290, Ap
411	12	23.1	363	4	US-09-252-991A-4960	Sequence 4960, Ap	484	12	23.1	513	4	US-09-252-991A-14193	Sequence 14193, A
412	12	23.1	363	4	US-09-252-991A-5444	Sequence 5444, Ap	485	12	23.1	516	4	US-09-252-991A-15293	Sequence 15293, A
413	12	23.1	366	4	US-09-252-991A-9508	Sequence 9508, Ap	486	12	23.1	519	4	US-09-252-991A-1363	Sequence 1363, Ap
414	12	23.1	387	4	US-09-252-991A-1461	Sequence 1461, Ap	487	12	23.1	519	4	US-09-252-991A-1363	Sequence 1363, Ap
415	12	23.1	387	4	US-09-252-991A-4967	Sequence 4967, Ap	488	12	23.1	519	4	US-09-252-991A-7319	Sequence 7319, Ap
416	12	23.1	393	4	US-09-252-991A-8006	Sequence 8006, Ap	489	12	23.1	519	4	US-09-252-991A-8226	Sequence 8226, Ap
417	12	23.1	399	1	US-07-985-692-1	Sequence 1, Appl1	490	12	23.1	531	4	US-09-252-991A-15586	Sequence 15586, A
418	12	23.1	399	1	US-08-155-331-1	Sequence 1, Appl1	491	12	23.1	531	4	US-09-199-637A-378	Sequence 378, Ap
419	12	23.1	399	1	US-08-424-022-1	Sequence 1, Appl1	492	12	23.1	543	4	US-09-252-991A-1325	Sequence 1325, Ap
420	12	23.1	399	1	US-08-424-017B-1	Sequence 1, Appl1	493	12	23.1	549	4	US-09-252-991A-1126	Sequence 1126, Ap
421	12	23.1	399	5	PCT-US93-11696-1	Sequence 1, Appl1	494	12	23.1	554	4	US-09-252-991A-6268	Sequence 6268, Ap
422	12	23.1	401	1	US-08-466-033-214	Sequence 214, Ap	495	12	23.1	558	4	US-09-252-991A-4967	Sequence 4967, Ap
423	12	23.1	401	1	US-08-444-733-214	Sequence 214, Ap	496	12	23.1	558	4	US-09-252-991A-6919	Sequence 6919, Ap
424	12	23.1	401	2	US-08-464-134-214	Sequence 214, Ap	497	12	23.1	561	4	US-09-252-991A-7502	Sequence 7502, Ap
425	12	23.1	401	2	US-08-461-361-214	Sequence 214, Ap	498	12	23.1	561	4	US-09-252-991A-12293	Sequence 12293, A
426	12	23.1	401	2	US-08-485-910-214	Sequence 214, Ap	499	12	23.1	564	4	US-09-199-637A-286	Sequence 286, Ap
427	12	23.1	402	4	US-09-199-637A-380	Sequence 380, Ap	500	12	23.1	564	4	US-09-252-991A-4893	Sequence 4893, Ap
428	12	23.1	405	4	US-09-702-705-1485	Sequence 1485, Ap	501	12	23.1	567	4	US-09-252-991A-15637	Sequence 15637, A
429	12	23.1	405	4	US-09-736-457-1485	Sequence 1485, Ap	502	12	23.1	570	4	US-09-252-991A-8189	Sequence 8189, Ap
430	12	23.1	405	4	US-09-252-991A-5615	Sequence 5615, Ap	503	12	23.1	570	4	US-09-252-991A-485	Sequence 485, Ap
431	12	23.1	408	4	US-09-252-991A-8130	Sequence 8130, Ap	504	12	23.1	579	4	US-09-252-991A-15611	Sequence 15611, A
432	12	23.1	411	4	US-09-252-991A-16270	Sequence 16270, A	505	12	23.1	579	4	US-09-199-637A-8	Sequence 8, Appl1
433	12	23.1	433	3	US-08-905-323-187	Sequence 16077, Ap	506	12	23.1	582	4	US-09-252-991A-9991	Sequence 9991, Ap
434	12	23.1	436	4	US-09-252-991A-16077	Sequence 16077, A	507	12	23.1	582	4	US-09-252-991A-371	Sequence 371, Ap
435	12	23.1	439	4	US-09-252-991A-8801	Sequence 8801, Ap	508	12	23.1	590	3	US-09-252-991A-12596	Sequence 12596, A
436	12	23.1	442	4	US-09-252-991A-1615	Sequence 1615, Ap	509	12	23.1	597	4	US-08-998-416-208	Sequence 208, Ap
437	12	23.1	432	4	US-09-252-991A-4796	Sequence 4796, Ap	510	12	23.1	599	2	US-09-252-991A-11757	Sequence 11757, A
438	12	23.1	435	4	US-09-252-991A-758	Sequence 758, Ap	511	12	23.1	600	4	US-08-665-647-8	Sequence 8, Appl1
439	12	23.1	440	3	US-09-060-756-544	Sequence 544, Ap	512	12	23.1	602	3	US-09-252-991A-13432	Sequence 13432, A
440	12	23.1	440	4	US-09-670-314-544	Sequence 544, Ap	513	12	23.1	602	3	US-09-078-294-27	Sequence 27, Appl1
441	12	23.1	441	4	US-09-252-991A-1192	Sequence 1192, Ap	514	12	23.1	603	4	US-09-252-991A-4719	Sequence 4719, Ap
442	12	23.1	441	4	US-09-252-991A-3890	Sequence 3890, Ap	515	12	23.1	603	4	US-09-252-991A-5732	Sequence 5732, Ap
443	12	23.1	441	4	US-09-252-991A-15107	Sequence 15107, A	516	12	23.1	609	4	US-09-252-991A-3856	Sequence 3856, Ap
444	12	23.1	441	4	US-09-252-991A-16152	Sequence 16152, A	517	12	23.1	612	4	US-09-252-991A-4424	Sequence 4424, Ap
445	12	23.1	444	4	US-09-199-637A-384	Sequence 384, Ap	518	12	23.1	618	4	US-09-199-637A-55	Sequence 55, Appl1
446	12	23.1	444	4	US-09-107-532A-2467	Sequence 2467, Ap	519	12	23.1	621	4	US-09-221-017B-75	Sequence 75, Appl1
447	12	23.1	447	4	US-09-252-991A-1705	Sequence 1705, Ap	520	12	23.1	621	4	US-09-252-991A-3944	Sequence 3944, Ap
448	12	23.1	447	4	US-09-252-991A-6438	Sequence 6438, Ap	521	12	23.1	624	4	US-09-252-991A-7192	Sequence 7192, Ap
449	12	23.1	448	3	US-09-252-991A-10307	Sequence 10307, A	522	12	23.1	624	3	US-09-181-974-1	Sequence 1, Appl1
450	12	23.1	448	3	US-09-060-756-179	Sequence 179, Ap	523	12	23.1	627	3	US-08-518-950-1	Sequence 9, Appl1
451	12	23.1	448	4	US-09-670-314-179	Sequence 179, Ap	524	12	23.1	627	4	US-09-449-249-9	Sequence 9, Appl1
452	12	23.1	450	4	US-09-252-991A-11345	Sequence 11246, A	525	12	23.1	637	4	US-09-352-991A-6070	Sequence 6070, Ap
453	12	23.1	450	4	US-09-252-991A-12127	Sequence 12127, A	526	12	23.1	637	4	US-09-252-991A-11427	Sequence 11427, A
454	12	23.1	450	4	US-09-252-991A-15147	Sequence 15147, A	527	12	23.1	639	4	US-09-252-991A-6398	Sequence 6398, Ap
455	12	23.1	456	4	US-09-252-991A-1545	Sequence 1545, Ap	528	12	23.1	640	4	US-09-252-991A-10430	Sequence 10430, A
456	12	23.1	458	3	US-09-252-991A-11687	Sequence 11687, A	529	12	23.1	639	4	US-09-252-991A-10430	Sequence 10430, A
457	12	23.1	458	3	US-09-060-756-138	Sequence 138, Ap	530	12	23.1	642	4	US-09-252-991A-14738	Sequence 14738, A
458	12	23.1	458	3	US-09-670-314-138	Sequence 138, Ap	531	12	23.1	642	4	US-09-252-991A-9255	Sequence 9255, Ap
459	12	23.1	461	4	US-09-857-556A-21	Sequence 21, Appl1	532	12	23.1	645	4	US-09-252-991A-11468	Sequence 11468, A
460	12	23.1	462	4	US-09-252-991A-9374	Sequence 9374, Ap	533	12	23.1	645	4	US-09-252-991A-2856	Sequence 2856, Ap
461	12	23.1	462	4	US-09-252-991A-10454	Sequence 10454, A	534	12	23.1	645	4	US-09-252-991A-5596	Sequence 5596, Ap
462	12	23.1	462	4	US-09-252-991A-10454	Sequence 10454, A	535	12	23.1	650	4	US-09-252-991A-16112	Sequence 16112, A
463	12	23.1	465	4	US-09-252-991A-235	Sequence 235, Ap	536	12	23.1	650	4	US-09-321-017B-647	Sequence 647, Ap
464	12	23.1	465	4	US-09-252-991A-10129	Sequence 10129, A	537	12	23.1	650	4	US-09-252-991A-11549	Sequence 11549, A
465	12	23.1	474	3	US-09-060-756-86	Sequence 86, Appl1	538	12	23.1	660	4	US-09-199-637A-414	Sequence 414, Ap

C 539	12	23.1	663	4	US-09-252-991A-1065	Sequence 1065, Ap	C 612	12	23.1	876	4	US-09-252-991A-6233	Sequence 6233, Ap
C 540	12	23.1	663	4	US-09-252-991A-9138	Sequence 181, App	C 613	12	23.1	878	2	US-08-827-279-2	Sequence 2, App1
C 541	12	23.1	663	4	US-09-328-352-181	Sequence 1622, Ap	C 614	12	23.1	882	4	US-09-205-255-167	Sequence 167, App1
C 542	12	23.1	666	4	US-09-252-991A-1622	Sequence 1622, Ap	C 615	12	23.1	885	4	US-09-252-991A-8741	Sequence 8741, Ap
C 543	12	23.1	669	4	US-09-252-991A-3703	Sequence 13532, A	C 616	12	23.1	885	4	US-09-252-991A-15757	Sequence 15757, A
C 544	12	23.1	671	3	US-08-866-340-23	Sequence 23, App1	C 617	12	23.1	891	4	US-09-252-991A-939	Sequence 939, App
C 545	12	23.1	671	3	US-09-103-875-29	Sequence 29, App1	C 618	12	23.1	891	4	US-09-252-991A-967	Sequence 967, App
C 546	12	23.1	671	3	US-09-252-991A-15543	Sequence 15543, A	C 619	12	23.1	891	4	US-09-252-991A-5380	Sequence 5380, App
C 547	12	23.1	672	4	US-08-998-416-199	Sequence 199, App	C 620	12	23.1	891	4	US-09-252-991A-14166	Sequence 14166, A
C 548	12	23.1	684	3	US-08-998-416-672	Sequence 672, App1	C 621	12	23.1	897	1	US-08-454-720A-32	Sequence 32, App1
C 549	12	23.1	684	3	US-09-651-169A-44	Sequence 44, App1	C 622	12	23.1	900	4	US-09-252-991A-4831	Sequence 4831, App
C 550	12	23.1	689	3	US-08-998-416-713	Sequence 713, App	C 623	12	23.1	900	4	US-09-252-991A-165	Sequence 165, App
C 551	12	23.1	689	3	US-09-252-991A-5173	Sequence 5173, App	C 624	12	23.1	903	4	US-09-252-991A-110472	Sequence 10472, A
C 552	12	23.1	690	4	US-09-252-991A-5297	Sequence 5297, App	C 625	12	23.1	909	4	US-09-252-991A-11108	Sequence 11108, A
C 553	12	23.1	693	4	US-09-252-991A-697	Sequence 697, App	C 626	12	23.1	909	4	US-09-252-991A-329	Sequence 329, App
C 554	12	23.1	693	4	US-09-328-352-946	Sequence 946, App	C 627	12	23.1	911	4	US-09-252-991A-11391	Sequence 11391, A
C 555	12	23.1	699	2	US-08-824-577-2	Sequence 2, App1	C 628	12	23.1	915	4	US-09-443-041A-23	Sequence 23, App1
C 556	12	23.1	702	4	US-08-824-577-2	Sequence 11915, A	C 629	12	23.1	915	4	US-09-221-011B-631	Sequence 631, App
C 557	12	23.1	702	4	US-09-484-970B-112	Sequence 112, App	C 630	12	23.1	921	4	US-09-252-991A-15590	Sequence 15590, A
C 558	12	23.1	708	4	US-09-265-965-39	Sequence 39, App1	C 631	12	23.1	933	4	US-09-199-631A-10	Sequence 10, App1
C 559	12	23.1	710	4	US-09-535-008-47	Sequence 47, App1	C 632	12	23.1	936	4	US-09-252-991A-16286	Sequence 16286, A
C 560	12	23.1	717	4	US-09-252-991A-13138	Sequence 13138, A	C 633	12	23.1	939	4	US-09-252-991A-13394	Sequence 13394, A
C 561	12	23.1	717	4	US-09-252-991A-6816	Sequence 6816, App	C 634	12	23.1	948	4	US-09-252-991A-11445	Sequence 11445, A
C 562	12	23.1	723	2	US-08-618-911-1	Sequence 1, App1	C 635	12	23.1	954	4	US-09-252-991A-4954	Sequence 4954, App
C 563	12	23.1	723	2	US-08-618-911-1	Sequence 142, App	C 636	12	23.1	954	4	US-09-252-991A-1935	Sequence 1935, App
C 564	12	23.1	730	4	US-09-663-600A-142	Sequence 5366, App	C 637	12	23.1	954	4	US-09-252-991A-1761	Sequence 1761, App
C 565	12	23.1	735	4	US-09-252-991A-5366	Sequence 5366, App	C 638	12	23.1	955	4	US-08-735-965-1	Sequence 1, App1
C 566	12	23.1	735	4	US-09-252-991A-5802	Sequence 5802, App	C 639	12	23.1	955	2	US-09-105-057-1	Sequence 1, App1
C 567	12	23.1	744	4	US-09-252-991A-10328	Sequence 10328, App	C 640	12	23.1	955	3	US-09-304-214-1	Sequence 1, App1
C 568	12	23.1	747	4	US-09-252-991A-11489	Sequence 11489, A	C 641	12	23.1	955	4	US-09-399-911-1	Sequence 11, App1
C 569	12	23.1	747	4	US-09-252-991A-16428	Sequence 16428, A	C 642	12	23.1	955	4	US-09-298-731-11	Sequence 11, App1
C 570	12	23.1	750	3	US-09-167-717-3	Sequence 8043, App	C 643	12	23.1	957	4	US-09-252-991A-13712	Sequence 13712, A
C 571	12	23.1	753	4	US-09-252-991A-6043	Sequence 13060, A	C 644	12	23.1	963	4	US-09-252-991A-16261	Sequence 16261, A
C 572	12	23.1	762	4	US-09-252-991A-13060	Sequence 5185, App	C 645	12	23.1	969	4	US-09-252-991A-8706	Sequence 8706, App
C 573	12	23.1	768	4	US-09-252-991A-5185	Sequence 15297, A	C 646	12	23.1	972	4	US-09-252-991A-13601	Sequence 13601, A
C 574	12	23.1	768	4	US-09-252-991A-15343	Sequence 284, App	C 647	12	23.1	978	4	US-09-252-991A-15990	Sequence 15990, A
C 575	12	23.1	774	4	US-09-252-991A-15297	Sequence 3285, App	C 648	12	23.1	981	4	US-09-252-991A-8278	Sequence 8278, App
C 576	12	23.1	780	4	US-09-252-991A-3285	Sequence 3285, App	C 649	12	23.1	984	4	US-09-252-991A-15726	Sequence 15726, A
C 577	12	23.1	783	4	US-09-252-991A-12804	Sequence 12723, A	C 650	12	23.1	987	4	US-09-641-638-117	Sequence 427, App
C 578	12	23.1	786	4	US-09-252-991A-15343	Sequence 555, App	C 651	12	23.1	992	4	US-09-252-991A-13041	Sequence 13041, A
C 579	12	23.1	789	4	US-09-252-991A-555	Sequence 48, App1	C 652	12	23.1	996	4	US-09-221-011B-770	Sequence 770, App
C 580	12	23.1	791	4	US-09-663-600A-48	Sequence 107, App1	C 653	12	23.1	1002	4	US-09-252-991A-4945	Sequence 13455, A
C 581	12	23.1	796	4	US-09-1280-116-107	Sequence 56, App1	C 654	12	23.1	1005	4	US-09-252-991A-1907	Sequence 7907, App
C 582	12	23.1	798	4	US-09-199-631A-56	Sequence 15941, A	C 655	12	23.1	1014	4	US-09-252-991A-2224	Sequence 2224, App
C 583	12	23.1	801	4	US-09-252-991A-14692	Sequence 284, App	C 656	12	23.1	1017	4	US-09-252-991A-11684	Sequence 11684, A
C 584	12	23.1	819	4	US-09-199-631A-384	Sequence 4858, App	C 657	12	23.1	1020	4	US-09-252-991A-11774	Sequence 11774, A
C 585	12	23.1	819	4	US-09-252-991A-4858	Sequence 15630, A	C 658	12	23.1	1023	4	US-09-252-991A-8251	Sequence 8251, App
C 586	12	23.1	822	4	US-09-252-991A-15630	Sequence 8940, App	C 659	12	23.1	1026	4	US-09-252-991A-14075	Sequence 14075, A
C 587	12	23.1	824	4	US-09-252-991A-8940	Sequence 11668, A	C 660	12	23.1	1041	4	US-09-252-991A-14894	Sequence 14894, A
C 588	12	23.1	825	4	US-09-252-991A-11668	Sequence 576, App	C 661	12	23.1	1044	4	US-09-391-741A-19	Sequence 741A-19
C 589	12	23.1	837	4	US-09-252-991A-576	Sequence 430, App	C 662	12	23.1	1045	4	US-09-391-741A-29	Sequence 741A-29
C 590	12	23.1	837	4	US-09-252-991A-5273	Sequence 5273, App	C 663	12	23.1	1047	2	US-08-494-907-11	Sequence 907-11
C 591	12	23.1	840	4	US-09-252-991A-5273	Sequence 430, App	C 664	12	23.1	1050	3	PRT-US98-10866-11	Sequence 10866-11
C 592	12	23.1	840	4	US-09-620-312D-318	Sequence 318, App	C 665	12	23.1	1050	3	US-08-681-192-1	Sequence 192-1
C 593	12	23.1	846	4	US-09-252-991A-2292	Sequence 3292, App	C 666	12	23.1	1056	4	US-09-252-991A-8728	Sequence 8728, App
C 594	12	23.1	852	4	US-09-252-991A-859	Sequence 859, App	C 667	12	23.1	1059	3	US-09-102-204-2	Sequence 2, App1
C 595	12	23.1	855	4	US-09-252-991A-8597	Sequence 859, App	C 668	12	23.1	1062	4	US-09-252-991A-884	Sequence 884, App
C 596	12	23.1	855	4	US-09-252-991A-9597	Sequence 9597, App	C 669	12	23.1	1062	4	US-09-252-991A-253	Sequence 253, App
C 597	12	23.1	855	4	US-09-252-991A-1297	Sequence 1297, App	C 670	12	23.1	1070	4	US-09-443-041A-9	Sequence 9, App1
C 598	12	23.1	864	4	US-09-252-991A-1297	Sequence 5034, App	C 671	12	23.1	1071	4	US-09-328-352-3545	Sequence 3545, App
C 599	12	23.1	870	4	US-09-252-991A-5034	Sequence 1268, App	C 672	12	23.1	1074	4	US-09-170-496D-23	Sequence 23, App1
C 600	12	23.1	872	4	US-09-252-991A-1268	Sequence 10645, A	C 673	12	23.1	1080	4	US-09-170-496D-175	Sequence 175, App
C 601	12	23.1	872	4	US-09-620-312D-316	Sequence 316, App	C 674	12	23.1				
C 602	12	23.1	873	4	US-09-252-991A-6226	Sequence 6226, App	C 675	12	23.1				
C 603	12	23.1	873	4	US-09-252-991A-14978	Sequence 14978, A	C 676	12	23.1				
C 604	12	23.1	876	2	US-08-937-972-2	Sequence 2, App1	C 677	12	23.1				
C 605	12	23.1	876	2	US-08-937-972-2	Sequence 924, App	C 678	12	23.1				
C 606	12	23.1	876	2	US-08-937-972-2	Sequence 924, App	C 679	12	23.1				
C 607	12	23.1	876	2	US-08-937-972-2	Sequence 924, App	C 680	12	23.1				
C 608	12	23.1	876	2	US-08-937-972-2	Sequence 924, App	C 681	12	23.1				
C 609	12	23.1	876	2	US-08-937-972-2	Sequence 924, App	C 682	12	23.1				
C 610	12	23.1	876	2	US-08-937-972-2	Sequence 924, App	C 683	12	23.1				
C 611	12	23.1	876	2	US-08-937-972-2	Sequence 924, App	C 684	12	23.1				

C 685	12	23.1	1080	4	US-09-252-991A-9553	Sequence 9553, Ap
C 686	12	23.1	1083	2	US-08-319-866-1	Sequence 1, App1
C 687	12	23.1	1082	4	US-09-252-991A-14341	Sequence 14341, A
C 688	12	23.1	1092	4	US-09-252-991A-14759	Sequence 14759, A
C 689	12	23.1	1095	4	US-09-252-991A-10116	Sequence 10116, A
C 690	12	23.1	1101	4	US-09-252-991A-15510	Sequence 15510, A
C 691	12	23.1	1107	4	US-09-227-357-19	Sequence 19, App1
C 692	12	23.1	1110	4	US-09-252-991A-8883	Sequence 8883, Ap
C 693	12	23.1	1110	4	US-09-252-991A-11012	Sequence 11012, A
C 694	12	23.1	1122	4	US-09-252-991A-3683	Sequence 3683, Ap
C 695	12	23.1	1122	4	US-09-252-991A-12348	Sequence 12348, A
C 696	12	23.1	1122	4	US-09-252-991A-14057	Sequence 14057, A
C 697	12	23.1	1131	4	US-09-252-991A-6825	Sequence 6825, Ap
C 698	12	23.1	1131	4	US-09-252-991A-9699	Sequence 9699, Ap
C 699	12	23.1	1131	4	US-09-252-991A-9699	Sequence 9699, Ap
C 700	12	23.1	1140	4	US-09-252-991A-5332	Sequence 5332, Ap
C 701	12	23.1	1143	4	US-09-252-991A-5748	Sequence 5748, Ap
C 702	12	23.1	1145	4	US-09-221-0178-944	Sequence 944, App
C 703	12	23.1	1152	4	US-09-252-991A-1656	Sequence 1656, Ap
C 704	12	23.1	1155	4	US-09-163-951-15	Sequence 15, App1
C 705	12	23.1	1161	4	US-09-252-991A-11607	Sequence 11607, A
C 706	12	23.1	1164	4	US-09-252-991A-4249	Sequence 4249, Ap
C 707	12	23.1	1167	4	US-09-252-991A-5001	Sequence 5001, Ap
C 708	12	23.1	1167	4	US-09-252-991A-14278	Sequence 14278, A
C 709	12	23.1	1170	4	US-09-252-991A-15019	Sequence 15019, A
C 710	12	23.1	1173	4	US-09-252-991A-3942	Sequence 3942, Ap
C 711	12	23.1	1185	4	US-09-252-991A-3103	Sequence 3103, Ap
C 712	12	23.1	1185	4	US-09-252-991A-5635	Sequence 5635, Ap
C 713	12	23.1	1184	4	US-09-232-991A-1326	Sequence 13928, A
C 714	12	23.1	1194	3	US-08-860-3688-11	Sequence 11, App1
C 715	12	23.1	1194	3	US-08-860-3688-12	Sequence 12, App1
C 716	12	23.1	1194	3	US-08-860-3688-13	Sequence 13, App1
C 717	12	23.1	1194	4	US-09-506-2868-21	Sequence 21, App1
C 718	12	23.1	1194	4	US-09-506-2868-25	Sequence 25, App1
C 719	12	23.1	1194	4	US-09-762-8618-21	Sequence 21, App1
C 720	12	23.1	1194	4	US-09-762-8618-25	Sequence 25, App1
C 721	12	23.1	1200	4	US-09-252-991A-278	Sequence 278, App
C 722	12	23.1	1206	4	US-09-252-991A-3338	Sequence 3328, Ap
C 723	12	23.1	1206	4	US-09-252-991A-5101	Sequence 5101, Ap
C 724	12	23.1	1206	4	US-09-252-991A-6454	Sequence 6454, App
C 725	12	23.1	1206	3	US-08-860-519-7	Sequence 7, App1
C 726	12	23.1	1218	4	US-09-252-991A-9482	Sequence 9482, Ap
C 727	12	23.1	1218	4	US-09-252-991A-16453	Sequence 16453, A
C 728	12	23.1	1237	4	US-09-252-991A-5204	Sequence 5204, Ap
C 729	12	23.1	1237	4	US-09-252-991A-10523	Sequence 10523, A
C 730	12	23.1	1237	4	US-09-252-991A-13527	Sequence 13527, A
C 731	12	23.1	1237	4	US-09-252-991A-16560	Sequence 16560, A
C 732	12	23.1	1237	4	US-09-252-991A-16560	Sequence 16560, A
C 733	12	23.1	1232	4	US-09-506-2868-22	Sequence 22, App1
C 734	12	23.1	1232	4	US-09-506-2868-23	Sequence 23, App1
C 735	12	23.1	1232	4	US-09-762-8618-22	Sequence 22, App1
C 736	12	23.1	1232	4	US-09-762-8618-23	Sequence 23, App1
C 737	12	23.1	1232	4	US-09-506-2868-19	Sequence 19, App1
C 738	12	23.1	1233	4	US-09-252-991A-14035	Sequence 14035, A
C 739	12	23.1	1233	4	US-09-762-8618-19	Sequence 19, App1
C 740	12	23.1	1233	4	US-09-252-991A-6096	Sequence 6096, Ap
C 741	12	23.1	1240	4	US-09-163-951-14	Sequence 14, App1
C 742	12	23.1	1248	4	US-09-252-991A-10189	Sequence 10189, A
C 743	12	23.1	1254	4	US-09-252-991A-305	Sequence 305, App
C 744	12	23.1	1257	4	US-09-252-991A-1067	Sequence 1027, Ap
C 745	12	23.1	1257	4	US-09-252-991A-15817	Sequence 15817, A
C 746	12	23.1	1257	4	US-09-328-352-921	Sequence 921, App
C 747	12	23.1	1260	4	US-08-305-1728-3	Sequence 3, App1
C 748	12	23.1	1260	4	US-09-252-991A-6674	Sequence 6674, Ap
C 749	12	23.1	1260	4	US-09-252-991A-14282	Sequence 14282, A
C 750	12	23.1	1263	1	US-08-532-8288-11	Sequence 11, App1
C 751	12	23.1	1263	1	US-08-532-8288-12	Sequence 12, App1
C 752	12	23.1	1263	4	US-09-252-991A-8847	Sequence 8847, Ap
C 753	12	23.1	1269	4	US-09-252-991A-3750	Sequence 3750, App
C 754	12	23.1	1270	4	US-09-016-434-756	Sequence 756, App
C 755	12	23.1	1272	4	US-09-252-991A-1552	Sequence 1552, Ap
C 756	12	23.1	1280	4	US-09-096-7768-4	Sequence 4, App1
C 757	12	23.1	1280	4	US-09-923-922-4	Sequence 4, App1
C 758	12	23.1	1281	4	US-09-252-991A-10951	Sequence 10951, A
C 759	12	23.1	1281	4	US-09-252-991A-11603	Sequence 11603, A
C 760	12	23.1	1282	2	US-08-878-989-12	Sequence 12, App1
C 761	12	23.1	1282	3	US-09-572-796-12	Sequence 12, App1
C 762	12	23.1	1282	4	US-09-016-434-953	Sequence 953, App
C 763	12	23.1	1287	4	US-09-054-680-1	Sequence 1, App1
C 764	12	23.1	1287	4	US-09-054-680-3	Sequence 3, App1
C 765	12	23.1	1287	4	US-09-252-991A-3628	Sequence 3628, Ap
C 766	12	23.1	1289	2	US-08-344-833-1	Sequence 1, App1
C 767	12	23.1	1290	4	US-09-252-991A-270	Sequence 270, App
C 768	12	23.1	1290	4	US-09-252-991A-9349	Sequence 9349, Ap
C 769	12	23.1	1293	4	US-09-252-991A-16545	Sequence 16545, A
C 770	12	23.1	1293	4	US-09-252-991A-16090	Sequence 16090, A
C 771	12	23.1	1301	2	US-08-641-3140-1	Sequence 1, App1
C 772	12	23.1	1305	4	US-09-252-991A-8472	Sequence 8472, Ap
C 773	12	23.1	1305	4	US-09-252-991A-8472	Sequence 8472, Ap
C 774	12	23.1	1305	4	US-09-252-991A-11762	Sequence 11762, A
C 775	12	23.1	1308	4	US-09-252-991A-1391	Sequence 1391, A
C 776	12	23.1	1311	4	US-09-252-991A-1391	Sequence 1391, A
C 777	12	23.1	1314	4	US-09-252-991A-11984	Sequence 11984, A
C 778	12	23.1	1314	4	US-09-252-991A-7168	Sequence 7168, Ap
C 779	12	23.1	1317	4	US-09-252-991A-11605	Sequence 11605, A
C 780	12	23.1	1317	4	US-09-252-991A-779	Sequence 779, App
C 781	12	23.1	1317	4	US-09-252-991A-9952	Sequence 9952, Ap
C 782	12	23.1	1317	4	US-09-252-991A-2299	Sequence 2299, Ap
C 783	12	23.1	1329	1	US-08-716-301-1	Sequence 1, App1
C 784	12	23.1	1335	2	US-08-743-6378-182	Sequence 182, App
C 785	12	23.1	1335	4	US-09-252-991A-5087	Sequence 5087, App
C 786	12	23.1	1335	4	US-09-252-991A-16196	Sequence 16196, A
C 787	12	23.1	1338	4	US-09-252-991A-7783	Sequence 7783, Ap
C 788	12	23.1	1338	4	US-09-252-991A-14144	Sequence 14144, A
C 789	12	23.1	1338	3	US-09-188-930-227	Sequence 227, App
C 790	12	23.1	1348	4	US-09-312-2830-227	Sequence 227, App
C 791	12	23.1	1352	4	US-09-016-434-1233	Sequence 1233, Ap
C 792	12	23.1	1352	5	PCT-US92-02091-7	Sequence 7, App1
C 793	12	23.1	1353	4	US-09-252-991A-15113	Sequence 15113, A
C 794	12	23.1	1356	4	US-09-252-991A-630	Sequence 630, App
C 795	12	23.1	1356	4	US-09-252-991A-11031	Sequence 11031, A
C 796	12	23.1	1356	4	US-09-252-991A-14188	Sequence 11081, A
C 797	12	23.1	1359	3	US-09-188-930-34	Sequence 34, App1
C 798	12	23.1	1359	3	US-09-312-2830-34	Sequence 34, App1
C 799	12	23.1	1362	4	US-09-470-5121-6	Sequence 6, App1
C 800	12	23.1	1365	4	US-09-252-991A-882	Sequence 882, App
C 801	12	23.1	1365	4	US-09-252-991A-8265	Sequence 8265, Ap
C 802	12	23.1	1365	4	US-09-280-116-120	Sequence 120, App
C 803	12	23.1	1368	1	US-08-278-7292-24	Sequence 24, App1
C 804	12	23.1	1368	1	US-08-155-433A-24	Sequence 24, App1
C 805	12	23.1	1368	1	US-08-406-672-24	Sequence 24, App1
C 806	12	23.1	1368	1	US-08-643-563A-24	Sequence 24, App1
C 807	12	23.1	1368	1	US-08-643-563A-24	Sequence 24, App1
C 808	12	23.1	1368	1	US-08-462-623-24	Sequence 24, App1
C 809	12	23.1	1368	1	US-08-451-953A-24	Sequence 24, App1
C 810	12	23.1	1368	2	US-08-445-668A-24	Sequence 24, App1
C 811	12	23.1	1368	2	US-08-461-397A-24	Sequence 24, App1
C 812	12	23.1	1368	2	US-08-912-088-24	Sequence 24, App1
C 813	12	23.1	1368	3	US-08-373-730A-24	Sequence 24, App1
C 814	12	23.1	1368	3	US-08-445-467-24	Sequence 24, App1
C 815	12	23.1	1368	3	US-08-480-515A-24	Sequence 24, App1
C 816	12	23.1	1368	3	US-08-271-556A-1	Sequence 1, App1
C 817	12	23.1	1368	3	US-09-170-936-24	Sequence 24, App1
C 818	12	23.1	1368	4	US-08-461-113-24	Sequence 24, App1
C 819	12	23.1	1368	4	US-08-456-033-24	Sequence 24, App1
C 820	12	23.1	1368	4	US-08-445-467-24	Sequence 24, App1
C 821	12	23.1	1368	4	US-09-252-991A-1810	Sequence 1810, Ap
C 822	12	23.1	1368	4	US-09-252-991A-5583	Sequence 5583, Ap
C 823	12	23.1	1368	4	US-09-252-991A-8408	Sequence 8408, Ap
C 824	12	23.1	1368	4	US-09-252-991A-16040	Sequence 16040, A
C 825	12	23.1	1368	5	PCT-US93-07199-24	Sequence 24, App1
C 826	12	23.1	1368	5	PCT-US93-07231-24	Sequence 24, App1
C 827	12	23.1	1368	5	PCT-US93-08742-24	Sequence 24, App1
C 828	12	23.1	1368	5	PCT-US93-08808-24	Sequence 24, App1
C 829	12	23.1	1368	5	PCT-US93-08808-24	Sequence 24, App1
C 830	12	23.1	1368	5	PCT-US93-08808-24	Sequence 24, App1

C 831	12	23.1	1368	5	PCT-US93-08885-24	Sequence 24, Appl1	C 904	12	23.1	1540	4	US-09-298-731-9	Sequence 9, Appl1
C 832	12	23.1	1371	2	US-08-910-731-1	Sequence 1, Appl1	C 905	12	23.1	1542	4	US-09-252-991A-6947	Sequence 647, Ap
C 833	12	23.1	1371	2	US-08-910-731-7	Sequence 7, Appl1	C 906	12	23.1	1542	4	US-09-252-991A-7625	Sequence 7625, Ap
C 834	12	23.1	1371	2	US-08-795-395-1	Sequence 1, Appl1	C 907	12	23.1	1548	2	US-08-762-106-5	Sequence 5, Appl1
C 835	12	23.1	1371	4	US-09-252-991A-5005	Sequence 5005, Ap	C 908	12	23.1	1548	3	US-09-320-774-5	Sequence 5, Appl1
C 836	12	23.1	1374	4	US-09-252-991A-881	Sequence 881, App	C 909	12	23.1	1548	4	US-09-252-991A-3653	Sequence 3653, Ap
C 837	12	23.1	1389	4	US-09-252-991A-5133	Sequence 5133, App	C 910	12	23.1	1548	4	US-09-252-991A-8413	Sequence 8413, Ap
C 838	12	23.1	1395	4	US-09-252-991A-9264	Sequence 9264, Ap	C 911	12	23.1	1550	2	US-08-609-443B-17	Sequence 17, Appl
C 839	12	23.1	1402	1	US-08-447-965A-1	Sequence 1, Appl1	C 912	12	23.1	1550	4	US-08-851-896-17	Sequence 15327, A
C 840	12	23.1	1404	1	US-07-984-044A-1	Sequence 1, Appl1	C 913	12	23.1	1551	4	US-09-252-991A-15327	Sequence 16227, A
C 841	12	23.1	1404	1	US-08-458-393-1	Sequence 1, Appl1	C 914	12	23.1	1554	4	US-09-352-991A-16227	Sequence 7, Appl1
C 842	12	23.1	1404	4	US-09-252-991A-9209	Sequence 9209, Ap	C 915	12	23.1	1556	3	US-09-334-601-7	Sequence 2, Appl1
C 843	12	23.1	1404	4	US-09-252-991A-12291	Sequence 12291, A	C 916	12	23.1	1560	2	US-08-709-979A-2	Sequence 767, Ap
C 844	12	23.1	1404	4	US-09-252-991A-14213	Sequence 14213, A	C 917	12	23.1	1560	4	US-09-252-991A-7367	Sequence 11778, A
C 845	12	23.1	1405	2	US-08-035-558-3	Sequence 3, Appl1	C 918	12	23.1	1560	4	US-09-352-991A-1778	Sequence 12332, A
C 846	12	23.1	1405	2	US-08-682-847-5	Sequence 5, Appl1	C 919	12	23.1	1563	4	US-09-352-991A-13332	Sequence 9220, Ap
C 847	12	23.1	1405	3	US-09-063-676-1	Sequence 1, Appl1	C 920	12	23.1	1566	4	US-09-252-991A-9220	Sequence 5425, Ap
C 848	12	23.1	1407	4	US-09-252-991A-637	Sequence 637, App	C 921	12	23.1	1569	4	US-09-252-991A-5425	Sequence 10972, A
C 849	12	23.1	1416	4	US-09-252-991A-1100	Sequence 1100, Ap	C 922	12	23.1	1578	4	US-09-252-991A-10972	Sequence 15375, A
C 850	12	23.1	1419	1	US-08-103-739B-1	Sequence 1, Appl1	C 923	12	23.1	1581	2	US-09-252-991A-15375	Sequence 6, Appl1
C 851	12	23.1	1419	1	US-08-474-404-1	Sequence 1, Appl1	C 924	12	23.1	1581	2	US-08-762-106-6	Sequence 246, App
C 852	12	23.1	1419	2	US-08-485-845-1	Sequence 1, Appl1	C 925	12	23.1	1581	3	US-09-320-774-6	Sequence 1, Appl1
C 853	12	23.1	1419	2	US-08-482-714-1	Sequence 1, Appl1	C 926	12	23.1	1590	4	US-09-252-991A-246	Sequence 5938, Ap
C 854	12	23.1	1419	3	US-09-211-416-1	Sequence 1, Appl1	C 927	12	23.1	1594	2	US-08-955-713-1	Sequence 11089, A
C 855	12	23.1	1419	3	US-09-059-958-1	Sequence 1, Appl1	C 928	12	23.1	1596	4	US-09-252-991A-5938	Sequence 5468, Ap
C 856	12	23.1	1419	4	US-09-252-991A-8433	Sequence 8433, Ap	C 929	12	23.1	1599	4	US-09-252-991A-3468	Sequence 15402, A
C 857	12	23.1	1428	4	US-09-252-991A-8370	Sequence 8370, Ap	C 930	12	23.1	1599	4	US-09-252-991A-15402	Sequence 286, Ap
C 858	12	23.1	1434	4	US-09-252-991A-3351	Sequence 3351, Ap	C 931	12	23.1	1601	4	US-09-252-991A-2386	Sequence 15438, A
C 859	12	23.1	1434	4	US-09-252-991A-384	Sequence 384, App	C 932	12	23.1	1611	4	US-09-252-991A-15438	Sequence 112, App
C 860	12	23.1	1449	4	US-09-252-991A-3558	Sequence 3558, Ap	C 933	12	23.1	1614	4	US-08-818-112-12	Sequence 107, App
C 861	12	23.1	1449	4	US-09-252-991A-7573	Sequence 7573, Ap	C 934	12	23.1	1616	4	US-08-818-111-107	Sequence 112, App
C 862	12	23.1	1449	4	US-09-252-991A-14399	Sequence 14399, A	C 935	12	23.1	1616	4	US-09-056-556-117	Sequence 107, App
C 863	12	23.1	1452	4	US-09-252-991A-3387	Sequence 3387, Ap	C 936	12	23.1	1616	4	US-09-072-556-102	Sequence 112, App
C 864	12	23.1	1457	4	US-09-620-112D-408	Sequence 408, App	C 937	12	23.1	1616	4	US-09-072-556-102	Sequence 107, App
C 865	12	23.1	1458	4	US-09-352-991A-5093	Sequence 5093, Ap	C 938	12	23.1	1617	4	US-09-252-991A-7508	Sequence 7508, Ap
C 866	12	23.1	1459	2	US-08-824-707-1	Sequence 1, Appl1	C 939	12	23.1	1620	4	US-09-252-991A-2743	Sequence 2743, Ap
C 867	12	23.1	1464	4	US-09-339-159B-13	Sequence 16136, A	C 940	12	23.1	1620	4	US-09-252-991A-1683	Sequence 14683, A
C 868	12	23.1	1464	4	US-09-339-159B-13	Sequence 16136, A	C 941	12	23.1	1626	4	US-09-252-991A-1683	Sequence 16549, A
C 869	12	23.1	1470	4	US-09-252-991A-896	Sequence 896, App	C 942	12	23.1	1638	4	US-09-352-991A-13378	Sequence 12378, A
C 870	12	23.1	1470	4	US-09-252-991A-3239	Sequence 3239, App	C 943	12	23.1	1638	4	US-09-352-991A-13378	Sequence 15572, A
C 871	12	23.1	1476	4	US-09-252-991A-15187	Sequence 15187, A	C 944	12	23.1	1639	4	US-09-620-312D-317	Sequence 317, App
C 872	12	23.1	1487	4	US-09-252-991A-6794	Sequence 6794, App	C 945	12	23.1	1643	1	US-08-532-828B-1	Sequence 1, Appl1
C 873	12	23.1	1487	4	US-09-252-991A-6794	Sequence 6794, App	C 946	12	23.1	1643	1	US-08-532-828B-2	Sequence 2, Appl1
C 874	12	23.1	1488	4	US-09-252-991A-421	Sequence 421, App	C 947	12	23.1	1643	1	US-08-532-828B-7	Sequence 7, Appl1
C 875	12	23.1	1488	4	US-09-252-991A-5798	Sequence 5798, App	C 948	12	23.1	1643	1	US-08-532-828B-8	Sequence 8, Appl1
C 876	12	23.1	1491	4	US-09-252-991A-786	Sequence 786, App	C 949	12	23.1	1643	1	US-08-532-828B-9	Sequence 9, Appl1
C 877	12	23.1	1494	4	US-09-252-991A-15431	Sequence 15431, A	C 950	12	23.1	1643	1	US-08-532-828B-10	Sequence 10, Appl1
C 878	12	23.1	1500	4	US-09-252-991A-9248	Sequence 9248, Ap	C 951	12	23.1	1643	1	US-08-700-359-7	Sequence 7, Appl1
C 879	12	23.1	1503	4	US-09-252-991A-6192	Sequence 6192, Ap	C 952	12	23.1	1643	1	US-08-700-359-8	Sequence 8, Appl1
C 880	12	23.1	1506	4	US-09-252-991A-15902	Sequence 15902, A	C 953	12	23.1	1643	1	US-08-700-359-9	Sequence 9, Appl1
C 881	12	23.1	1509	4	US-09-252-991A-6930	Sequence 6930, Ap	C 954	12	23.1	1643	1	US-08-674-168-21	Sequence 21, Appl1
C 882	12	23.1	1510	1	US-08-361-920-26	Sequence 26, Appl1	C 955	12	23.1	1643	2	US-08-596-366-7	Sequence 5, Appl1
C 883	12	23.1	1510	1	US-08-479-939-26	Sequence 26, Appl1	C 956	12	23.1	1643	2	US-08-596-366-7	Sequence 7, Appl1
C 884	12	23.1	1510	1	US-08-483-432-26	Sequence 26, Appl1	C 957	12	23.1	1643	2	US-08-967-104-5	Sequence 5, Appl1
C 885	12	23.1	1511	2	US-08-809-763-3	Sequence 3, Appl1	C 958	12	23.1	1643	2	US-08-967-104-7	Sequence 7, Appl1
C 886	12	23.1	1511	3	US-08-956-253-3	Sequence 3, Appl1	C 959	12	23.1	1643	3	US-08-985-908-3	Sequence 3, Appl1
C 887	12	23.1	1512	4	US-09-199-637A-171	Sequence 171, App	C 960	12	23.1	1643	3	US-08-985-908-4	Sequence 4, Appl1
C 888	12	23.1	1512	4	US-09-252-991A-1138	Sequence 1138, App	C 961	12	23.1	1643	3	US-08-985-908-6	Sequence 6, Appl1
C 889	12	23.1	1518	4	US-08-576-0635-21	Sequence 21, Appl	C 962	12	23.1	1643	3	US-08-852-730-12	Sequence 12, Appl
C 890	12	23.1	1518	4	US-09-252-991A-2196	Sequence 2196, Ap	C 963	12	23.1	1643	3	US-08-852-730-13	Sequence 13, Appl
C 891	12	23.1	1518	4	US-09-252-991A-12883	Sequence 12883, A	C 964	12	23.1	1643	3	US-08-852-730-15	Sequence 15, Appl
C 892	12	23.1	1518	4	US-09-658-772-1	Sequence 1, Appl1	C 965	12	23.1	1643	3	US-08-852-730-15	Sequence 15, Appl
C 893	12	23.1	1521	4	US-09-252-991A-2479	Sequence 2479, Ap	C 966	12	23.1	1643	3	US-08-985-916-3	Sequence 3, Appl1
C 894	12	23.1	1524	2	US-08-716-942-24	Sequence 24, Appl	C 967	12	23.1	1643	3	US-08-985-916-4	Sequence 4, Appl1
C 895	12	23.1	1524	3	US-08-840-767-3	Sequence 24, Appl1	C 968	12	23.1	1645	2	US-08-985-916-6	Sequence 6, Appl1
C 896	12	23.1	1524	4	US-09-130-337A-24	Sequence 24, Appl1	C 969	12	23.1	1647	2	US-08-039-297B-1	Sequence 9, Appl1
C 897	12	23.1	1527	4	US-09-252-991A-7133	Sequence 7133, Ap	C 970	12	23.1	1647	2	US-08-198-446B-9	Sequence 9, Appl1
C 898	12	23.1	1528	4	US-09-016-434-1448	Sequence 1448, Ap	C 971	12	23.1	1647	2	US-08-870-693-9	Sequence 9, Appl1
C 899	12	23.1	1531	4	US-08-630-915A-21	Sequence 21, Appl	C 972	12	23.1	1647	4	US-09-207-388-9	Sequence 9, Appl1
C 900	12	23.1	1533	4	US-09-252-991A-16513	Sequence 16513, A	C 973	12	23.1	1650	4	US-09-252-991A-8097	Sequence 8097, Ap
C 901	12	23.1	1534	4	US-09-399-913-7	Sequence 7, Appl1	C 974	12	23.1	1659	1	US-08-333-358-7	Sequence 7, Appl1
C 902	12	23.1	1534	4	US-09-298-731-7	Sequence 7, Appl1	C 975	12	23.1	1659	1	US-08-463-694-7	Sequence 7, Appl1
C 903	12	23.1	1540	4	US-09-399-913-9	Sequence 9, Appl1	C 976	12	23.1	1665	4	US-08-694-501-7	Sequence 7, Appl1
												US-09-252-991A-4934	Sequence 4934, Ap

```
C 977 12 23.1 1668 4 US-09-252-991A-12163 Sequence 12163, A
C 978 12 23.1 1674 4 US-09-252-991A-5894 Sequence 5894, Ap
C 979 12 23.1 1677 4 US-09-199-637A-276 Sequence 276, App
980 12 23.1 1677 4 US-09-252-991A-6747 Sequence 6747, Ap
981 12 23.1 1677 4 US-09-252-991A-7476 Sequence 7476, Ap
C 982 12 23.1 1683 4 US-09-252-991A-6977 Sequence 6977, Ap
C 983 12 23.1 1683 4 US-09-252-991A-11588 Sequence 11588, A
984 12 23.1 1692 4 US-09-252-991A-978 Sequence 978, App
C 985 12 23.1 1692 4 US-09-252-991A-14477 Sequence 14477, Ap
C 986 12 23.1 1692 4 US-09-252-991A-14777 Sequence 14777, A
C 987 12 23.1 1698 4 US-09-252-991A-3551 Sequence 3551, Ap
C 988 12 23.1 1700 3 US-08-860-519-5 Sequence 5, Appl1
C 989 12 23.1 1700 3 US-08-860-519-6 Sequence 6, Appl1
990 12 23.1 1701 4 US-09-252-991A-7927 Sequence 7927, Ap
991 12 23.1 1707 4 US-09-207-388-10 Sequence 10, Appl
992 12 23.1 1710 4 US-09-252-991A-9057 Sequence 9057, Ap
C 993 12 23.1 1723 1 US-08-241-766-1 Sequence 1, Appl1
C 994 12 23.1 1723 1 US-08-241-766-2 Sequence 2, Appl1
C 995 12 23.1 1724 3 US-09-385-259-1 Sequence 1, Appl1
C 996 12 23.1 1724 4 US-09-645-370-1 Sequence 1, Appl1
997 12 23.1 1728 4 US-09-252-991A-616 Sequence 616, App
998 12 23.1 1734 4 US-09-252-991A-3996 Sequence 3996, Ap
999 12 23.1 1734 4 US-09-252-991A-9061 Sequence 9061, Ap
1000 12 23.1 1736 4 US-09-162-524-2 Sequence 2, Appl1
```

ALIGNMENTS

```
RESULT 1
US-09-252-991A-13272/c
; Sequence 13272, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13272
; LENGTH: 732
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13272

Query Match 38.5%; Score 20; DB 4; Length 732;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 CTCGTCGTCGCGGCGCTCGC 47
Db 381 CTCGTCGTCGCGGCGCTCGC 362

RESULT 2
US-09-252-991A-13107/c
; Sequence 13107, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13107
; LENGTH: 732
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13107
```

```
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13107
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13107
```

```
Query Match 38.5%; Score 20; DB 4; Length 1056;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 28 CTCGTCGTCGCGGCGCTCGC 47
Db 634 CTCGTCGTCGCGGCGCTCGC 615
```

```
RESULT 3
US-09-252-991A-3647/c
; Sequence 3647, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3647
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3647
```

```
Query Match 30.8%; Score 16; DB 4; Length 810;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 33 CGCTGCCGCGGCGCTCGC 48
Db 286 CGCTGCCGCGGCGCTCGC 271
```

```
RESULT 4
US-09-252-991A-3602/c
; Sequence 3602, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3602
; LENGTH: 879
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3602
```

```
Query Match 30.8%; Score 16; DB 4; Length 879;
Best Local Similarity 100.0%; Pred. No. 7.9;
```

```
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 CGCTGCGGCGCTGCC 48
Db 338 CGCTGCGGCGCTGCC 323

RESULT 5
US-09-252-991A-3859
; Sequence 3859, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3859
; LENGTH: 1632
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3859

Query Match 30.8%; Score 16; DB 4; Length 1632;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 CGCTGCGGCGCTGCC 48
Db 1353 CGCTGCGGCGCTGCC 1368

RESULT 6
US-09-252-991A-2993/C
; Sequence 2993, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2993
; LENGTH: 2673
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2993

Query Match 30.8%; Score 16; DB 4; Length 2673;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 GCTCTGTCGTCGCCG 40
Db 133 GCTCTGTCGTCGCCG 118

RESULT 7
US-09-252-991A-2686
; Sequence 2686, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2686
; LENGTH: 2715
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2686

Query Match 30.8%; Score 16; DB 4; Length 2715;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 GCTCTGTCGTCGCCG 40
Db 2652 GCTCTGTCGTCGCCG 2667

RESULT 8
US-09-252-991A-2875
; Sequence 2875, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2875
; LENGTH: 2814
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2875

Query Match 30.8%; Score 16; DB 4; Length 2814;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 GCTCTGTCGTCGCCG 40
Db 2710 GCTCTGTCGTCGCCG 2725

RESULT 9
US-09-072-596-259/C
; Sequence 259, Application US/09072596
; Patent No. 6458356
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS
```

NUMBER OF SEQUENCES: 350
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,596
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-6031
FAX: (206) 622-4900
INFORMATION FOR SEQ ID NO: 259:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-072-596-259

Query Match 28.8%; Score 15; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 GCTCGCTCGCGCCT 44
|||||
Db 26 GCTCGCTCGCGCCT 12

RESULT 10
US-09-252-991A-15918
Sequence 15918, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 15918
LENGTH: 438
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15918

Query Match 28.8%; Score 15; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 34 GCTCGCGCTCGCC 48
|||||
Db 136 GCTCGCGCTCGCC 150

RESULT 11
US-09-060-756-60
Sequence 60, Application US/09060756
Patent No. 6183957
GENERAL INFORMATION:
APPLICANT: Cole, Stewart
APPLICANT: Buchrieser-Brosch, Roland
APPLICANT: Gordon, Stephen
APPLICANT: Billault, Alain
TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
FILE REFERENCE: 3495-0169
CURRENT APPLICATION NUMBER: US/09/060,756
CURRENT FILING DATE: 1998-04-16
NUMBER OF SEQ ID NOS: 743
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 60
LENGTH: 448
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
NAME/KEY: unsure
LOCATION: (various positions within the sequence)
OTHER INFORMATION: applicants are uncertain of bases designated as "n"
US-09-060-756-60

Query Match 28.8%; Score 15; DB 3; Length 448;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TTCTCTCAGCTCTC 21
|||||
Db 352 TTCTCTCAGCTCTC 366

RESULT 12
US-09-670-314-60
Sequence 60, Application US/09670314
Patent No. 6492506
GENERAL INFORMATION:
APPLICANT: Cole, Stewart
APPLICANT: Buchrieser-Brosch, Roland
APPLICANT: Gordon, Stephen
APPLICANT: Billault, Alain
TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
FILE REFERENCE: 3495-0169
CURRENT APPLICATION NUMBER: US/09/670,314
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 09/060,756
PRIOR FILING DATE: 1998-04-16
NUMBER OF SEQ ID NOS: 743
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 60
LENGTH: 448
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
NAME/KEY: unsure
LOCATION: (various positions within the sequence)
OTHER INFORMATION: applicants are uncertain of bases designated as "n"
US-09-670-314-60

Query Match 28.8%; Score 15; DB 4; Length 448;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 TTCTCTCAGCTCTC 21
|||||
Db 352 TTCTCTCAGCTCTC 366


```
RESULT 13
US-09-252-991A-15888
; Sequence 15888, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15888
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15888

Query Match      28.8%; Score 15; DB 4; Length 762;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      34 GCTGCCGGCCTCGCC 48
Db      485 GCTGCCGGCCTCGCC 499

RESULT 14
US-09-252-991A-15774/C
; Sequence 15774, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15774
; LENGTH: 885
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15774

Query Match      28.8%; Score 15; DB 4; Length 885;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      34 GCTGCCGGCCTCGCC 48
Db      274 GCTGCCGGCCTCGCC 260

RESULT 15
US-09-252-991A-15804/C
; Sequence 15804, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
```

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15804
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15804

Query Match      28.8%; Score 15; DB 4; Length 1068;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      34 GCTGCCGGCCTCGCC 48
Db      495 GCTGCCGGCCTCGCC 481

RESULT 16
US-09-252-991A-10788/C
; Sequence 10788, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10788
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10788

Query Match      28.8%; Score 15; DB 4; Length 1119;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      34 GCTGCCGGCCTCGCC 48
Db      685 GCTGCCGGCCTCGCC 671

RESULT 17
US-09-252-991A-10603/C
; Sequence 10603, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10603
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10603
```


Query Match 28.8%; Score 15; DB 4; Length 1146;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 GCTGCGCGCTCGCC 48
|||||
Db 735 GCTGCGCGCTCGCC 721

RESULT 18

US-09-620-312D-168/C
; Sequence 168, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunxui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radote T.
; TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 168
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)..(548)
US-09-620-312D-168

Query Match 28.8%; Score 15; DB 4; Length 1197;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 CTCATCGCTCTGCTC 33
|||||
Db 100 CTCATCGCTCTGCTC 86

RESULT 19

US-09-252-991A-10250
; Sequence 10250, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10250
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10250

Query Match 28.8%; Score 15; DB 4; Length 1260;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 GCTGCGCGCTCGCC 48
|||||
Db 406 GCTGCGCGCTCGCC 420

RESULT 20

US-09-252-991A-11745/C
; Sequence 11745, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11745
; LENGTH: 1317
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11745

Query Match 28.8%; Score 15; DB 4; Length 1317;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 TCGTCGCTGCCGCC 43
|||||
Db 629 TCGTCGCTGCCGCC 615

RESULT 21

US-09-252-991A-10368
; Sequence 10368, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10368
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10368

Query Match 28.8%; Score 15; DB 4; Length 1620;

Best Local Similarity 100.0%; Pred. No. 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 GCTGCGGCGCTGCC 48
Db 1266 GCTGCGGCGCTGCC 1280

RESULT 22
US-09-252-991A-11887/c
; Sequence 11887, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11887
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11887

Query Match 28.8%; Score 15; DB 4; Length 1692;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 TCCTCGCTCGCGGCC 43
Db 273 TCCTCGCTCGCGGCC 259

RESULT 23
US-09-252-991A-113
; Sequence 113, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 113
; LENGTH: 2553
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-113

Query Match 28.8%; Score 15; DB 4; Length 2553;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 CGTCGCTCGCGGCT 44
Db 105 CGTCGCTCGCGGCT 119

RESULT 24
US-09-252-991A-108/c
; Sequence 108, Application US/09252991A

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 108
; LENGTH: 2664
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-108

Query Match 28.8%; Score 15; DB 4; Length 2664;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 30 CGTCGCTCGCGGCT 44
Db 2449 CGTCGCTCGCGGCT 2435

RESULT 25
US-09-586-719-11
; Sequence 11, Application US/09586719
; Patent No. 6294345
; GENERAL INFORMATION:
; APPLICANT: Zheng, Feng
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Bauer, Michael W.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; FILE REFERENCE: PB/5-10909A
; CURRENT APPLICATION NUMBER: US/09/586,719
; PRIOR FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 4198
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-586-719-11

Query Match 28.8%; Score 15; DB 3; Length 4198;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TCTCTCAGCCTCTCA 22
Db 958 TCTCTCAGCCTCTCA 972

RESULT 26
5183745-1/c
; Patent No. 5183745
; APPLICANT: DANCHIN, ANTOINE; GLASER, PHILIPPE; KRIN, EVELYN;
; BARZU, OCTAVIEN; LANDANT, DANIEL; ULLMAN, AGNES
; TITLE OF INVENTION: ADENYL CYCLASE DERIVATIVES AND THEIR
; BIOLOGICAL USES
; NUMBER OF SEQUENCES: 13
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,541
; FILING DATE: 25-OCT-1989
; SEQ ID NO 1:
; LENGTH: 4649
5183745-1

Query Match 28.8%; Score 15; DB 6; Length 4649;

Best Local Similarity 100.0%; Pred. No. 23;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 GCTGCCGCGCTCGCC 48
|||||
Db 4350 GCTGCCGCGCTCGCC 4336

RESULT 27

US-08-669-785-3/c
; Sequence 3, Application US/08669785
; Patent No. 6309648
; GENERAL INFORMATION:
; APPLICANT: Betsou, Fotini
; APPLICANT: Betsou, Peter
; APPLICANT: Guiso, Nicole
; TITLE OF INVENTION: Protective Epitopes Of Adenyl
; TITLE OF INVENTION: Cyclase-Haemolysin(AC-Hly), Their Application To
; TITLE OF INVENTION: The Treatment Or To The Prevention Of Bordetella Infections.
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3115
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,785
; FILING DATE: 27-JUN-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0072-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5118 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..5115
; OTHER INFORMATION: /note= "AMINO ACID SEQUENCE
; OTHER INFORMATION: CORRESPONDING TO THE NUCLEOTIDE SEQUENCE OF THE GENE
; OTHER INFORMATION: CODING FOR THE B. Bronchiseptica AC-Hly"
US-08-669-785-3

Query Match 28.8%; Score 15; DB 4; Length 5118;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 GCTGCCGCGCTCGCC 48
|||||
Db 4164 GCTGCCGCGCTCGCC 4150

RESULT 28
US-08-669-785-1/c
; Sequence 1, Application US/08669785
; Patent No. 6309648
; GENERAL INFORMATION:
; APPLICANT: Betsou, Fotini

; APPLICANT: Sebo, Peter
; APPLICANT: Guiso, Nicole
; TITLE OF INVENTION: Protective Epitopes Of Adenyl
; TITLE OF INVENTION: Cyclase-Haemolysin(AC-Hly), Their Application To
; TITLE OF INVENTION: The Treatment Or To The Prevention Of Bordetella Infections.
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3115
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,785
; FILING DATE: 27-JUN-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0072-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6441 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 979..6096
; OTHER INFORMATION: /note= "AMINO ACID SEQUENCE
; OTHER INFORMATION: CORRESPONDING TO THE NUCLEOTIDE SEQUENCE OF THE GENE
; OTHER INFORMATION: CODING FOR THE B. Pertussis AC-Hly"
US-08-669-785-1

Query Match 28.8%; Score 15; DB 4; Length 6441;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 GCTGCCGCGCTCGCC 48
|||||
Db 5145 GCTGCCGCGCTCGCC 5131

RESULT 29
US-08-669-785-5/c
; Patent No. 5183745
; APPLICANT: DANCHIN, ANTOINE; GLASER, PHILIPPE; KRIN, EVELYN;
; BARZU, OCTAVIEN; LADANT, DANIEL; ULLMAN, AGNES
; TITLE OF INVENTION: ADENYL CYCLASE DERIVATIVES AND THEIR
; BIOLOGICAL USES
; NUMBER OF SEQUENCES: 13
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,541
; FILING DATE: 25-OCT-1989
; SEQ ID NO: 5:
; LENGTH: 6443
US-08-669-785-5

Query Match 28.8%; Score 15; DB 6; Length 6443;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 GCTGCCGCGCTCGCC 48

```
Db          5147 GCTGCCGCTGCC 5133
|||||
RESULT 30
US-09-453-702B-242/C
; Sequence 242, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blatner, Frederick R.
;           Burland, Valerie
;           Perna, Nicole T.
;           Plunkett, Guy
;           Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Quarles & Brady
; STREET: 1 South Plunkney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 242:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31880
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 242:
US-09-453-702B-242

Query Match      28.8%; Score 15; DB 4; Length 31880;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      23 TCGCTCTGTCGCTG 37
Db      26165 TCGCTCTGTCGCTG 26151

RESULT 31
US-09-922-445-1
; Sequence 1, Application US/09922445
; Patent No. 6528268
; GENERAL INFORMATION:
; APPLICANT: Anderson, Maria K.
;           Berglund, Lars G. T.
;           Berglund, Rikard H.
;           Applicant: Adam, Gail I. R.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR DETECTION OF HEART FAILURE
; FILE REFERENCE: GGI2605
; CURRENT APPLICATION NUMBER: US/09/922,445

; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 1
; LENGTH: 38653
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(26156)
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (24801)..(24801)
; OTHER INFORMATION: nucleotide 24801 is a single nucleotide polymorphism which can t
; NAME/KEY: misc_feature
; LOCATION: (24941)..(24941)
; OTHER INFORMATION: nucleotide 24941 is a single nucleotide polymorphism which can t
; OTHER INFORMATION: T or C
; NAME/KEY: exon
; LOCATION: (26157)..(26252)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (26253)..(26401)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (26402)..(26543)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (26544)..(27024)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (27025)..(27178)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (27179)..(30519)
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (27645)..(27645)
; OTHER INFORMATION: nucleotide 27645 is a single nucleotide polymorphism which can t
; OTHER INFORMATION: C or G
; NAME/KEY: exon
; LOCATION: (30520)..(30681)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (30682)..(30894)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (30895)..(31027)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (31028)..(31747)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (31748)..(31841)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (31842)..(32400)
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (32163)..(32163)
; OTHER INFORMATION: nucleotide 32163 is a single nucleotide polymorphism which can b
; OTHER INFORMATION: A or C
; NAME/KEY: exon
; LOCATION: (32401)..(32528)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (32529)..(33414)
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (32614)..(32614)
; OTHER INFORMATION: nucleotide 32614 is a single nucleotide polymorphism which can b
; OTHER INFORMATION: A or G
```

```

; NAME/KEY: exon
; LOCATION: (33415)..(33597)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (33598)..(34314)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (34315)..(34588)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (34589)..(36404)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (36405)..(36523)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (36524)..(38341)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (38342)..(38653)
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank/AC004923
; DATABASE ENTRY DATE: 1999-12-21
; RELEVANT RESIDUES: (1)..(38653)
US-09-922-445-1

```

```

Query Match          28.8%; Score 15; DB 4; Length 38653;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      15 CGCTTCATCGCTCT 29
Db      8343 CGCTTCATCGCTCT 8357

```

```

RESULT 32
US-08-804-227C-7
; Sequence 7, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rosteck, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44377 base pairs
; TYPE: nucleic acid
;

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..14002
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830
US-08-804-227C-7

```

```

Query Match          28.8%; Score 15; DB 2; Length 44377;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      34 GCTGCCGCGCTCGCC 48
Db      17373 GCTGCCGCGCTCGCC 17387

```

```

RESULT 33
US-08-804-198-1
; Sequence 1, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rosteck, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P9113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3885
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..14002
;

```

```
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830
; US-08-804-198-1
```

```
Query Match          28.8%; Score 15; DB 2; Length 44377;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      34 GCTGCCGCGCTCGCC 48
          |||||
Db       17373 GCTGCCGCGCTCGCC 17387
```

```
RESULT 34
US-09-335-409-1/c
; Sequence 1, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goetlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-335-409-1
```

```
Query Match          28.8%; Score 15; DB 3; Length 68750;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      32 TCGCTGCCGCGCTCG 46
          |||||
Db       6467 TCGCTGCCGCGCTCG 6453
```

```
RESULT 35
US-09-568-102-1/c
; Sequence 1, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goetlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
```

```
;
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-568-102-1
```

```
Query Match          28.8%; Score 15; DB 4; Length 68750;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      32 TCGCTGCCGCGCTCG 46
          |||||
Db       6467 TCGCTGCCGCGCTCG 6453
```

```
RESULT 36
US-09-567-969-1/c
; Sequence 1, Application US/09567969
; Patent No. 6355457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goetlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,969
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-567-969-1
```

```
Query Match          28.8%; Score 15; DB 4; Length 68750;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      32 TCGCTGCCGCGCTCG 46
          |||||
Db       6467 TCGCTGCCGCGCTCG 6453
```

```
RESULT 37
US-09-568-480-1/c
; Sequence 1, Application US/09568480
; Patent No. 6355458
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goetlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,480
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
```

US-09-568-480-1

Query Match 28.8%; Score 15; DB 4; Length 68750;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 TCGCTGCCGCGCCTCG 46
|||||
DB 6467 TCGCTGCCGCGCCTCG 6453

RESULT 38

US-09-568-486-1/c
; Sequence 1, Application US/09568486
; Patent No. 6355459
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,486
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-568-486-1

Query Match 28.8%; Score 15; DB 4; Length 68750;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 TCGCTGCCGCGCCTCG 46
|||||
DB 6467 TCGCTGCCGCGCCTCG 6453

RESULT 39

US-09-568-472-1/c
; Sequence 1, Application US/09568472
; Patent No. 6358719
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,472
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-568-472-1

Query Match 28.8%; Score 15; DB 4; Length 68750;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 TCGCTGCCGCGCCTCG 46
|||||
DB 6467 TCGCTGCCGCGCCTCG 6453

RESULT 40

US-09-567-899-1/c
; Sequence 1, Application US/09567899
; Patent No. 6383787
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,899
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-567-899-1

Query Match 28.8%; Score 15; DB 4; Length 68750;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 TCGCTGCCGCGCCTCG 46
|||||
DB 6467 TCGCTGCCGCGCCTCG 6453

Search completed: November 13, 2003, 10:59:54
Job time : 79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2003, 10:26:42 ; Search time 214 Seconds

(without alignments)
794.143 Million cell updates/sec

Title: US-10-081-935-1

Perfect score: 52
Sequence: 1 atgcagctctctcacgctct.....cgctgcgagctcgcgcagtrg 52

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2169961 seqs, 1634102185 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications, NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	17	32.7	1392	12	US-10-286-606-11
C 2	17	32.7	1392	14	US-10-126-279-11
C 3	16	30.8	574	11	US-09-918-995-25977
C 4	16	30.8	1653	9	US-09-897-214-7
C 5	16	30.8	1872	9	US-09-768-826-21
C 6	16	30.8	2451	12	US-10-093-463-189
C 7	16	30.8	2563	12	US-10-093-463-191
C 8	16	30.8	2610	12	US-10-246-310-1
C 9	16	30.8	3201	12	US-10-093-463-187
C 10	16	30.8	3947	10	US-09-925-300-638
C 11	16	30.8	3953	14	US-10-106-698-1931
C 12	15	28.8	264	9	US-09-294-093B-4207
C 13	15	28.8	264	10	US-09-878-574-15763
C 14	15	28.8	294	14	US-10-156-761-711
C 15	15	28.8	336	10	US-09-878-574-1259
C 16	15	28.8	357	9	US-09-764-870-129

C 17	15	28.8	357	14	US-10-125-540-129	Sequence 129, App
C 18	15	28.8	423	12	US-10-084-843-264	Sequence 264, App
C 19	15	28.8	423	12	US-10-193-002-259	Sequence 259, App
C 20	15	28.8	481	12	US-10-259-678-60	Sequence 60, App
C 21	15	28.8	481	9	US-09-864-761-6099	Sequence 6099, App
C 22	15	28.8	603	12	US-10-027-633-115584	Sequence 115584, App
C 23	15	28.8	603	13	US-10-027-633-115584	Sequence 115584, App
C 24	15	28.8	1113	14	US-10-156-761-974	Sequence 974, App
C 25	15	28.8	1191	12	US-10-259-165-279	Sequence 279, App
C 26	15	28.8	1197	14	US-10-037-270-168	Sequence 168, App
C 27	15	28.8	1389	14	US-10-156-761-760	Sequence 760, App
C 28	15	28.8	1404	14	US-10-177-2293-416	Sequence 416, App
C 29	15	28.8	1479	14	US-10-156-761-4461	Sequence 4461, App
C 30	15	28.8	1554	14	US-10-128-714-7070	Sequence 7070, App
C 31	15	28.8	1778	14	US-10-128-714-6070	Sequence 6070, App
C 32	15	28.8	1932	10	US-09-801-368-235	Sequence 235, App
C 33	15	28.8	2000	10	US-09-938-842A-2978	Sequence 2978, App
C 34	15	28.8	2412	14	US-10-156-761-158	Sequence 158, App
C 35	15	28.8	2454	9	US-09-815-242-4138	Sequence 4138, App
C 36	15	28.8	2672	11	US-09-374-046A-125	Sequence 125, App
C 37	15	28.8	2790	10	US-09-954-456-518	Sequence 518, App
C 38	15	28.8	2790	10	US-09-967-768A-182	Sequence 182, App
C 39	15	28.8	3140	14	US-10-128-714-70	Sequence 70, App
C 40	15	28.8	3323	11	US-09-822-846-197	Sequence 197, App
C 41	15	28.8	3778	14	US-10-128-714-5070	Sequence 5070, App
C 42	15	28.8	3963	12	US-10-161-051-168	Sequence 168, App
C 43	15	28.8	4463	12	US-10-214-350-3	Sequence 3, App
C 44	15	28.8	6059	12	US-10-240-452-43	Sequence 43, App
C 45	15	28.8	6442	10	US-09-950-335A-11	Sequence 11, App
C 46	15	28.8	10431	14	US-10-198-846-11715	Sequence 11715, App
C 47	15	28.8	10438	12	US-09-814-351-19864	Sequence 19864, App
C 48	15	28.8	11447	14	US-10-270-333-187	Sequence 187, App
C 49	15	28.8	31880	14	US-10-114-170-242	Sequence 242, App
C 50	15	28.8	68750	13	US-10-014-171-1	Sequence 1, App
C 51	15	28.8	536165	11	US-09-939-964-1	Sequence 1, App
C 52	15	28.8	9026608	14	US-10-156-761-1	Sequence 1, App
C 53	15	28.8	9026608	14	US-10-156-761-1	Sequence 1, App
C 54	14	26.9	209	9	US-09-923-876-2874	Sequence 2874, App
C 55	14	26.9	234	9	US-09-876-889-36	Sequence 36, App
C 56	14	26.9	255	14	US-10-156-761-1349	Sequence 1349, App
C 57	14	26.9	310	11	US-09-866-050A-110	Sequence 110, App
C 58	14	26.9	310	14	US-10-152-661-1107	Sequence 1107, App
C 59	14	26.9	384	10	US-09-878-574-1507	Sequence 1507, App
C 60	14	26.9	420	14	US-10-156-761-1546	Sequence 1546, App
C 61	14	26.9	427	9	US-09-912-020-64	Sequence 64, App
C 62	14	26.9	432	13	US-10-033-159A-5	Sequence 5, App
C 63	14	26.9	468	11	US-09-918-995-24558	Sequence 24558, App
C 64	14	26.9	492	14	US-10-156-761-1986	Sequence 1986, App
C 65	14	26.9	498	12	US-10-027-632-180608	Sequence 180608, App
C 66	14	26.9	498	12	US-10-027-632-180608	Sequence 180608, App
C 67	14	26.9	498	13	US-10-027-632-180608	Sequence 180608, App
C 68	14	26.9	498	13	US-10-027-632-180608	Sequence 180608, App
C 69	14	26.9	515	12	US-10-399-140-4	Sequence 140, App
C 70	14	26.9	516	12	US-10-027-632-36803	Sequence 36803, App
C 71	14	26.9	516	12	US-10-027-632-80036	Sequence 80036, App
C 72	14	26.9	516	13	US-10-027-632-36803	Sequence 36803, App
C 73	14	26.9	516	13	US-10-027-632-36803	Sequence 36803, App
C 74	14	26.9	522	9	US-09-764-870-572	Sequence 572, App
C 75	14	26.9	522	9	US-09-764-870-572	Sequence 572, App
C 76	14	26.9	522	14	US-10-125-540-572	Sequence 572, App
C 77	14	26.9	522	14	US-10-125-540-573	Sequence 573, App
C 78	14	26.9	524	12	US-10-027-632-273527	Sequence 273527, App
C 79	14	26.9	524	12	US-10-027-632-273528	Sequence 273528, App
C 80	14	26.9	524	13	US-10-027-632-273527	Sequence 273527, App
C 81	14	26.9	524	13	US-10-027-632-273528	Sequence 273528, App
C 82	14	26.9	526	12	US-10-027-632-273529	Sequence 273529, App
C 83	14	26.9	526	13	US-10-027-632-273529	Sequence 273529, App
C 84	14	26.9	539	12	US-10-027-632-2735107	Sequence 2735107, App
C 85	14	26.9	539	13	US-10-027-632-2735107	Sequence 2735107, App
C 86	14	26.9	567	12	US-10-259-165-277	Sequence 277, App
C 87	14	26.9	573	12	US-10-259-165-271	Sequence 271, App
C 88	14	26.9	576	12	US-10-259-165-271	Sequence 271, App
C 89	14	26.9	612	10	US-09-938-842A-129	Sequence 129, App

C 90	14	26.9	621	12	US-10-027-632-132148	Sequence 132148,	C 163	14	26.9	60196	14	US-10-205-032-1	Sequence 1, Appl1
C 91	14	26.9	621	13	US-10-027-632-132148	Sequence 132148,	C 164	14	26.9	96194	12	US-10-329-079-34	Sequence 34, Appl1
C 92	14	26.9	663	9	US-09-770-149-454	Sequence 454, App	C 165	14	26.9	96649	11	US-09-956-712-10	Sequence 10, Appl1
C 93	14	26.9	769	12	US-10-027-632-22263	Sequence 22263, A	C 166	14	26.9	171936	12	US-10-025-966A-24	Sequence 24, Appl1
C 94	14	26.9	769	13	US-10-027-632-22263	Sequence 22263, A	C 167	14	26.9	171936	12	US-10-265-071-24	Sequence 24, Appl1
C 95	14	26.9	789	9	US-09-815-242-4032	Sequence 4032, Ap	C 168	14	26.9	2940917	12	US-10-027-632-174763	Sequence 174763,
C 96	14	26.9	804	10	US-09-998-582-1	Sequence 1, Appl1	C 169	14	26.9	2940917	12	US-10-027-632-174763	Sequence 174763,
C 97	14	26.9	816	14	US-10-156-761-5917	Sequence 5917, Ap	C 170	14	26.9	3186778	13	US-10-027-632-174961	Sequence 174961,
C 98	14	26.9	844	12	US-10-027-632-146101	Sequence 146102,	C 171	14	26.9	3186778	13	US-10-027-632-174961	Sequence 174961,
C 99	14	26.9	844	12	US-10-027-632-146102	Sequence 146102,	C 172	14	26.9	3186778	13	US-10-027-632-174961	Sequence 174961,
C 100	14	26.9	844	13	US-10-027-632-146102	Sequence 146102,	C 173	13	25.0	309400	10	US-09-738-626-1	Sequence 1, Appl1
C 101	14	26.9	844	13	US-10-027-632-146102	Sequence 146102,	C 174	13	25.0	20	9	US-09-752-639-72	Sequence 72, Appl1
C 102	14	26.9	849	9	US-09-770-445-630	Sequence 630, App	C 175	13	25.0	20	10	US-09-984-196-77	Sequence 29833, A
C 103	14	26.9	960	14	US-10-007-521-1	Sequence 1, Appl1	C 176	13	25.0	65	12	US-09-908-975-25823	Sequence 29833, A
C 104	14	26.9	983	12	US-10-027-632-323876	Sequence 323876,	C 177	13	25.0	122	10	US-09-920-300A-115	Sequence 315, App
C 105	14	26.9	983	12	US-10-027-632-323876	Sequence 323876,	C 178	13	25.0	122	10	US-09-998-598-1589	Sequence 1589, App
C 106	14	26.9	983	13	US-10-027-632-323876	Sequence 323876,	C 179	13	25.0	122	12	US-10-099-928-315	Sequence 315, App
C 107	14	26.9	983	13	US-10-027-632-323876	Sequence 323876,	C 180	13	25.0	122	13	US-10-033-528-315	Sequence 315, App
C 108	14	26.9	1034	12	US-10-214-529-5	Sequence 323433,	C 181	13	25.0	172	10	US-09-878-574-13265	Sequence 13265, A
C 109	14	26.9	1101	13	US-10-032-159A-1	Sequence 1, Appl1	C 182	13	25.0	178	10	US-09-878-574-13265	Sequence 13265, A
C 110	14	26.9	1122	14	US-10-156-761-4958	Sequence 4958, Ap	C 183	13	25.0	186	12	US-10-029-386-23561	Sequence 23561, A
C 111	14	26.9	1332	14	US-10-156-761-2193	Sequence 2193, Ap	C 184	13	25.0	189	10	US-09-983-965-3376	Sequence 3376, Ap
C 112	14	26.9	1401	14	US-10-156-761-6368	Sequence 6368, Ap	C 185	13	25.0	206	12	US-10-029-386-21128	Sequence 21128, A
C 113	14	26.9	1608	11	US-09-798-412-6	Sequence 6, Appl1	C 186	13	25.0	208	9	US-09-864-761-22408	Sequence 22408, A
C 114	14	26.9	1608	14	US-10-325-917-6	Sequence 6, Appl1	C 187	13	25.0	225	10	US-09-998-598-2367	Sequence 2267, App
C 115	14	26.9	1617	10	US-09-738-626-2432	Sequence 2432, Ap	C 188	13	25.0	226	10	US-09-998-598-1541	Sequence 1541, App
C 116	14	26.9	1698	10	US-09-938-842A-2638	Sequence 2638, Ap	C 189	13	25.0	246	14	US-10-102-524-158	Sequence 148, App
C 117	14	26.9	1725	13	US-10-001-843-74	Sequence 74, Appl1	C 190	13	25.0	249	9	US-09-294-0938-757	Sequence 757, App
C 118	14	26.9	1746	10	US-09-938-842A-1453	Sequence 1453, Ap	C 191	13	25.0	249	14	US-09-974-300-4033	Sequence 4033, Ap
C 119	14	26.9	1767	14	US-10-156-761-3727	Sequence 3727, Ap	C 192	13	25.0	268	12	US-10-259-165-659	Sequence 1, Appl1
C 120	14	26.9	1857	10	US-09-965-553-24	Sequence 24, Appl1	C 193	13	25.0	268	12	US-10-259-165-659	Sequence 669, App
C 121	14	26.9	1966	12	US-10-017-161-2369	Sequence 2369, Ap	C 194	13	25.0	272	9	US-09-294-0938-2718	Sequence 2718, Ap
C 122	14	26.9	1970	14	US-10-772-417-1	Sequence 1, Appl1	C 195	13	25.0	275	10	US-09-878-574-1332	Sequence 1933, Ap
C 123	14	26.9	2036	10	US-09-783-252-18	Sequence 18, Appl1	C 196	13	25.0	284	9	US-09-294-0938-1538	Sequence 1538, Ap
C 124	14	26.9	2098	11	US-09-798-412-4	Sequence 4, Appl1	C 197	13	25.0	286	10	US-09-878-574-14763	Sequence 14763, A
C 125	14	26.9	2098	14	US-10-325-917-4	Sequence 4, Appl1	C 198	13	25.0	288	9	US-09-294-0938-4668	Sequence 4668, Ap
C 126	14	26.9	2176	13	US-10-032-159A-19	Sequence 19, Appl1	C 199	13	25.0	296	9	US-09-294-0938-1444	Sequence 1344, Ap
C 127	14	26.9	2436	14	US-10-156-761-1420	Sequence 1420, Ap	C 200	13	25.0	333	10	US-09-938-626-154	Sequence 2838, Ap
C 128	14	26.9	2634	9	US-09-391-340-7	Sequence 7, Appl1	C 201	13	25.0	336	10	US-09-938-626-154	Sequence 2075, Ap
C 129	14	26.9	2634	10	US-09-948-369-7	Sequence 7, Appl1	C 202	13	25.0	341	9	US-09-770-791-843	Sequence 843, App
C 130	14	26.9	2799	9	US-09-902-627-2	Sequence 2, Appl1	C 203	13	25.0	343	10	US-09-878-574-2608	Sequence 2608, Ap
C 131	14	26.9	2844	12	US-09-814-353-22021	Sequence 22021, A	C 204	13	25.0	345	10	US-09-933-797-155	Sequence 155, App
C 132	14	26.9	2894	14	US-10-128-714-223	Sequence 223, App	C 205	13	25.0	349	10	US-09-983-965-5391	Sequence 5391, App
C 133	14	26.9	2902	9	US-09-877-804-1	Sequence 1, Appl1	C 206	13	25.0	355	10	US-09-764-877-530	Sequence 530, App
C 134	14	26.9	3031	12	US-10-027-632-115642	Sequence 115642,	C 207	13	25.0	357	9	US-09-864-761-23562	Sequence 23562, A
C 135	14	26.9	3031	13	US-10-027-632-115642	Sequence 115642,	C 208	13	25.0	381	12	US-10-029-386-23070	Sequence 23070, A
C 136	14	26.9	3053	14	US-10-128-714-523	Sequence 5233, Ap	C 209	13	25.0	382	11	US-09-918-995-19031	Sequence 19031, A
C 137	14	26.9	3169	14	US-10-128-714-423	Sequence 4233, App	C 210	13	25.0	393	10	US-09-960-335-5954	Sequence 5954, Ap
C 138	14	26.9	3269	14	US-10-128-714-5423	Sequence 5423, Ap	C 211	13	25.0	395	11	US-09-918-995-4484	Sequence 4484, Ap
C 139	14	26.9	3464	10	US-09-848-852A-2	Sequence 2, Appl1	C 212	13	25.0	399	10	US-09-918-995-4484	Sequence 4484, Ap
C 140	14	26.9	3871	9	US-09-729-674-1	Sequence 1, Appl1	C 213	13	25.0	401	10	US-09-983-965-3957	Sequence 3957, App
C 141	14	26.9	3897	14	US-10-156-761-4899	Sequence 4899, Ap	C 214	13	25.0	401	11	US-09-997-181-75	Sequence 75, Appl1
C 142	14	26.9	3953	11	US-09-764-891-7337	Sequence 7337, Ap	C 215	13	25.0	401	11	US-09-997-181-75	Sequence 75, Appl1
C 143	14	26.9	3979	12	US-10-387-388-4	Sequence 4, Appl1	C 216	13	25.0	407	10	US-09-960-335-11022	Sequence 12022, A
C 144	14	26.9	4521	13	US-10-002-600-93	Sequence 93, Appl1	C 217	13	25.0	410	10	US-09-960-335-7977	Sequence 7977, Ap
C 145	14	26.9	4810	10	US-09-885-329-5	Sequence 5, Appl1	C 218	13	25.0	414	11	US-09-918-995-11554	Sequence 17554, A
C 146	14	26.9	4810	12	US-10-387-388-3	Sequence 3, Appl1	C 219	13	25.0	418	10	US-09-960-335-7653	Sequence 7853, Ap
C 147	14	26.9	5163	12	US-10-387-388-7	Sequence 7, Appl1	C 220	13	25.0	422	10	US-09-783-559-8355	Sequence 8355, Ap
C 148	14	26.9	5509	12	US-10-387-388-1	Sequence 1, Appl1	C 221	13	25.0	424	11	US-09-918-995-4259	Sequence 4259, Ap
C 149	14	26.9	6208	12	US-10-311-455-168	Sequence 168, App	C 222	13	25.0	426	10	US-09-960-335-11139	Sequence 11139, A
C 150	14	26.9	6868	11	US-09-919-039-39	Sequence 39, Appl1	C 223	13	25.0	427	10	US-09-960-335-9461	Sequence 9461, Ap
C 151	14	26.9	6868	12	US-10-133-013-139	Sequence 139, App	C 224	13	25.0	427	12	US-10-027-632-181320	Sequence 181320,
C 152	14	26.9	6981	12	US-09-873-319-635	Sequence 636, App	C 225	13	25.0	427	13	US-10-027-632-181320	Sequence 181320,
C 153	14	26.9	6981	12	US-09-960-706-977	Sequence 977, App	C 226	13	25.0	428	10	US-09-867-701-4567	Sequence 4567, Ap
C 154	14	26.9	7319	12	US-10-240-965-149	Sequence 299, App	C 227	13	25.0	432	11	US-09-918-995-35806	Sequence 35806, A
C 155	14	26.9	7460	14	US-10-176-847-93	Sequence 149, App	C 228	13	25.0	439	12	US-10-259-678-98	Sequence 98, Appl1
C 156	14	26.9	7460	14	US-10-176-847-93	Sequence 93, Appl1	C 229	13	25.0	444	9	US-09-770-444-833	Sequence 833, App
C 157	14	26.9	7494	12	US-10-161-051-190	Sequence 190, App	C 230	13	25.0	444	12	US-10-027-632-42571	Sequence 42571, A
C 158	14	26.9	11238	14	US-10-305-033-15	Sequence 15, Appl1	C 231	13	25.0	444	12	US-10-027-632-42572	Sequence 42572, A
C 159	14	26.9	13894	12	US-10-315-515-121	Sequence 121, App	C 232	13	25.0	444	12	US-10-027-632-42573	Sequence 42573, A
C 160	14	26.9	13966	12	US-10-315-515-122	Sequence 122, App	C 233	13	25.0	444	13	US-10-027-632-42572	Sequence 42572, A
C 161	14	26.9	15798	12	US-10-329-079-46	Sequence 46, Appl1	C 234	13	25.0	444	13	US-10-027-632-42572	Sequence 42572, A
C 162	14	26.9	25801	12	US-10-181-319-13	Sequence 13, Appl1	C 235	13	25.0	444	13	US-10-027-632-42573	Sequence 42573, A

C 236	13	25.0	448	11	US-09-918-995-12325	Sequence 12325, A	309	13	25.0	611	13	US-10-027-632-240833	Sequence 240833, A
C 237	13	25.0	458	11	US-09-918-995-23102	Sequence 23102, A	310	13	25.0	612	10	US-09-974-300-5291	Sequence 5291, Ap
C 238	13	25.0	460	11	US-09-918-995-29406	Sequence 29406, A	311	13	25.0	618	12	US-10-329-079-64	Sequence 64, Appl
C 239	13	25.0	460	12	US-10-027-632-142477	Sequence 142477, A	312	13	25.0	624	12	US-10-027-632-208118	Sequence 208118, A
C 240	13	25.0	460	13	US-10-027-632-142477	Sequence 142477, A	313	13	25.0	624	12	US-10-027-632-208119	Sequence 208119, A
C 241	13	25.0	462	9	US-09-864-761-5641	Sequence 5641, Ap	314	13	25.0	624	13	US-10-027-632-208118	Sequence 208118, A
C 242	13	25.0	474	11	US-09-918-995-22407	Sequence 22407, A	315	13	25.0	625	12	US-10-027-632-208119	Sequence 208119, A
C 243	13	25.0	479	11	US-09-918-995-22233	Sequence 22233, A	316	13	25.0	625	12	US-10-027-632-95884	Sequence 95884, A
C 244	13	25.0	479	12	US-10-027-632-287831	Sequence 287831, A	317	13	25.0	625	12	US-10-027-632-95884	Sequence 95884, A
C 245	13	25.0	479	13	US-10-027-632-287831	Sequence 287831, A	318	13	25.0	631	12	US-10-027-632-95884	Sequence 95884, A
C 246	13	25.0	480	9	US-09-864-761-423	Sequence 423, App	319	13	25.0	631	12	US-10-027-632-48036	Sequence 48036, A
C 247	13	25.0	480	11	US-09-918-995-1297	Sequence 1297, Ap	320	13	25.0	631	13	US-10-027-632-48036	Sequence 48036, A
C 248	13	25.0	483	9	US-09-864-761-6498	Sequence 6498, Ap	321	13	25.0	631	13	US-10-027-632-48037	Sequence 48037, A
C 249	13	25.0	489	11	US-09-918-995-32772	Sequence 32772, A	322	13	25.0	637	12	US-10-027-632-237904	Sequence 237904, A
C 250	13	25.0	495	14	US-10-156-761-29322	Sequence 29322, A	323	13	25.0	637	12	US-10-027-632-237905	Sequence 237905, A
C 251	13	25.0	500	11	US-09-918-995-1734	Sequence 1734, Ap	324	13	25.0	637	13	US-10-027-632-237904	Sequence 237904, A
C 252	13	25.0	501	10	US-09-796-692-3217	Sequence 3217, Ap	325	13	25.0	637	13	US-10-027-632-237905	Sequence 237905, A
C 253	13	25.0	501	14	US-10-040-862-3217	Sequence 3217, Ap	326	13	25.0	640	12	US-10-181-447A-60	Sequence 60, Appl
C 254	13	25.0	503	12	US-10-029-386-9036	Sequence 9036, Ap	327	13	25.0	650	12	US-10-181-447A-60	Sequence 60, Appl
C 255	13	25.0	504	12	US-10-027-632-135579	Sequence 135579, A	328	13	25.0	650	12	US-10-027-632-104429	Sequence 104429, A
C 256	13	25.0	504	12	US-10-027-632-135580	Sequence 135580, A	329	13	25.0	650	12	US-10-027-632-335520	Sequence 335520, A
C 257	13	25.0	504	13	US-10-027-632-135579	Sequence 135579, A	330	13	25.0	650	13	US-10-027-632-104429	Sequence 104429, A
C 258	13	25.0	504	13	US-10-027-632-135580	Sequence 135580, A	331	13	25.0	651	14	US-10-027-632-335520	Sequence 335520, A
C 259	13	25.0	507	12	US-10-027-632-109162	Sequence 109162, A	332	13	25.0	654	12	US-10-156-761-5660	Sequence 5660, Ap
C 260	13	25.0	507	12	US-10-027-632-109163	Sequence 109163, A	333	13	25.0	654	12	US-10-027-632-238748	Sequence 238748, A
C 261	13	25.0	507	12	US-10-027-632-180387	Sequence 180387, A	334	13	25.0	659	9	US-10-027-632-238748	Sequence 238748, A
C 262	13	25.0	507	12	US-10-027-632-300147	Sequence 300147, A	335	13	25.0	668	12	US-09-956-004-113	Sequence 113, App
C 263	13	25.0	507	12	US-10-027-632-300148	Sequence 300147, A	336	13	25.0	668	12	US-10-027-632-11165	Sequence 11165, A
C 264	13	25.0	507	13	US-10-027-632-109162	Sequence 109162, A	337	13	25.0	668	12	US-10-027-632-11167	Sequence 11167, A
C 265	13	25.0	507	13	US-10-027-632-109163	Sequence 109163, A	338	13	25.0	668	13	US-10-027-632-11165	Sequence 11165, A
C 266	13	25.0	507	13	US-10-027-632-180387	Sequence 180387, A	339	13	25.0	668	13	US-10-027-632-11166	Sequence 11166, A
C 267	13	25.0	507	13	US-10-027-632-300147	Sequence 300147, A	340	13	25.0	668	13	US-10-027-632-11167	Sequence 11167, A
C 268	13	25.0	507	13	US-10-027-632-300148	Sequence 300148, A	341	13	25.0	669	9	US-09-770-149-427	Sequence 427, App
C 269	13	25.0	513	10	US-09-833-381-972	Sequence 972, App	342	13	25.0	671	12	US-10-027-632-110507	Sequence 110507, A
C 270	13	25.0	522	12	US-10-027-632-275858	Sequence 275858, A	343	13	25.0	671	13	US-10-027-632-110507	Sequence 110507, A
C 271	13	25.0	522	13	US-10-027-632-275858	Sequence 275858, A	344	13	25.0	685	12	US-10-027-632-20685	Sequence 20685, A
C 272	13	25.0	527	12	US-10-027-632-33436	Sequence 33436, A	345	13	25.0	685	13	US-10-027-632-20685	Sequence 20685, A
C 273	13	25.0	527	13	US-10-027-632-33436	Sequence 33436, A	346	13	25.0	694	12	US-10-027-632-17732	Sequence 17732, A
C 274	13	25.0	528	11	US-09-918-995-23561	Sequence 23561, A	347	13	25.0	699	10	US-09-974-300-1200	Sequence 1200, Ap
C 275	13	25.0	534	14	US-10-156-761-5941	Sequence 5941, A	348	13	25.0	735	12	US-09-938-842A-918	Sequence 918, Ap
C 276	13	25.0	535	12	US-10-029-386-9861	Sequence 9861, Ap	349	13	25.0	735	12	US-10-027-632-133852	Sequence 133852, A
C 277	13	25.0	539	10	US-09-867-701-5697	Sequence 5697, Ap	350	13	25.0	735	12	US-10-027-632-133853	Sequence 133853, A
C 278	13	25.0	543	9	US-09-864-761-6829	Sequence 6829, Ap	351	13	25.0	735	12	US-10-027-632-147681	Sequence 147681, A
C 279	13	25.0	543	10	US-09-887-576-786	Sequence 786, App	352	13	25.0	735	12	US-10-027-632-133852	Sequence 133852, A
C 280	13	25.0	550	12	US-09-814-353-6069	Sequence 6069, Ap	353	13	25.0	735	13	US-10-027-632-133853	Sequence 133853, A
C 281	13	25.0	550	12	US-09-814-353-13348	Sequence 12348, A	354	13	25.0	736	12	US-10-027-632-147681	Sequence 147681, A
C 282	13	25.0	555	12	US-09-814-353-13348	Sequence 12348, A	355	13	25.0	736	12	US-10-027-632-12738	Sequence 12738, A
C 283	13	25.0	557	9	US-09-815-343-466	Sequence 9370, Ap	356	13	25.0	738	13	US-10-027-632-12738	Sequence 12738, A
C 284	13	25.0	557	10	US-09-938-842A-4261	Sequence 4261, Ap	357	13	25.0	738	10	US-09-738-636-2293	Sequence 2293, Ap
C 285	13	25.0	560	10	US-09-887-576-714	Sequence 714, App	358	13	25.0	751	12	US-10-027-632-158553	Sequence 158553, A
C 286	13	25.0	563	12	US-10-027-632-277127	Sequence 277127, A	359	13	25.0	751	13	US-10-027-632-158553	Sequence 158553, A
C 287	13	25.0	563	12	US-10-027-632-277127	Sequence 277127, A	360	13	25.0	764	9	US-09-925-297-269	Sequence 269, App
C 288	13	25.0	563	13	US-10-027-632-277127	Sequence 277127, A	361	13	25.0	765	12	US-10-027-632-22760	Sequence 22760, A
C 289	13	25.0	565	11	US-10-027-632-277128	Sequence 277128, A	362	13	25.0	775	13	US-10-027-632-22760	Sequence 22760, A
C 290	13	25.0	565	11	US-09-918-995-188	Sequence 188, App	363	13	25.0	780	10	US-09-996-634-70	Sequence 634-70
C 291	13	25.0	573	12	US-10-027-632-281808	Sequence 281808, A	364	13	25.0	780	11	US-09-997-162-70	Sequence 162-70
C 292	13	25.0	573	13	US-10-027-632-281808	Sequence 281808, A	365	13	25.0	780	11	US-09-997-162-70	Sequence 162-70
C 293	13	25.0	580	14	US-10-066-543-795	Sequence 795, App	366	13	25.0	789	9	US-09-997-161-70	Sequence 161-70
C 294	13	25.0	585	12	US-10-029-386-1674	Sequence 12674, A	367	13	25.0	792	11	US-09-866-060A-471	Sequence 471, App
C 295	13	25.0	586	12	US-10-027-632-321576	Sequence 321576, A	368	13	25.0	792	14	US-10-152-661-471	Sequence 471, App
C 296	13	25.0	588	12	US-10-029-386-7428	Sequence 321576, A	369	13	25.0	798	14	US-10-152-661-471	Sequence 471, App
C 297	13	25.0	588	12	US-10-029-386-7428	Sequence 7428, Ap	370	13	25.0	807	14	US-10-152-661-471	Sequence 471, App
C 298	13	25.0	590	11	US-09-918-995-13200	Sequence 12300, A	371	13	25.0	811	12	US-10-027-632-155593	Sequence 155593, A
C 299	13	25.0	590	12	US-10-027-632-132457	Sequence 132457, A	372	13	25.0	822	12	US-10-027-632-155593	Sequence 155593, A
C 300	13	25.0	590	13	US-10-027-632-132457	Sequence 132457, A	373	13	25.0	822	12	US-10-027-632-155593	Sequence 155593, A
C 301	13	25.0	593	12	US-10-027-632-104894	Sequence 104894, A	374	13	25.0	846	10	US-10-160-764-6	Sequence 6, Appl1
C 302	13	25.0	593	13	US-10-027-632-104894	Sequence 104894, A	375	13	25.0	846	10	US-10-160-764-6	Sequence 6, Appl1
C 303	13	25.0	594	12	US-10-029-386-5007	Sequence 5007, Ap	376	13	25.0	892	12	US-09-938-842A-262	Sequence 262, App
C 304	13	25.0	598	12	US-10-029-386-9558	Sequence 9558, Ap	377	13	25.0	903	12	US-10-029-386-22904	Sequence 22904, A
C 305	13	25.0	599	9	US-09-925-301-262	Sequence 262, App	378	13	25.0	903	14	US-10-251-385-471	Sequence 249, App
C 306	13	25.0	611	12	US-10-027-632-223826	Sequence 223826, A	379	13	25.0	903	14	US-10-251-385-471	Sequence 271, App
C 307	13	25.0	611	12	US-10-027-632-240833	Sequence 240833, A	380	13	25.0	903	14	US-10-106-698-2002	Sequence 2002, App
C 308	13	25.0	611	13	US-10-027-632-223826	Sequence 223826, A	381	13	25.0	903	14	US-10-225-567A-474	Sequence 474, App

C 382	13	25.0	911	9	US-09-864-761-17715	Sequence 17715, A	C 455	13	25.0	1035	9	US-09-815-242-4004	Sequence 4004, Ap
C 383	13	25.0	912	14	US-10-156-761-6808	Sequence 6808, Ap	C 456	13	25.0	1037	9	US-09-844-468-1	Sequence 1, Appl1
C 384	13	25.0	913	9	US-09-770-445-375	Sequence 375, App	C 457	13	25.0	1039	14	US-10-156-761-5718	Sequence 5718, Ap
C 385	13	25.0	914	12	US-10-259-165-603	Sequence 603, App	C 458	13	25.0	1067	10	US-09-917-800A-1529	Sequence 1529, Ap
C 386	13	25.0	954	14	US-10-156-761-1715	Sequence 1715, Ap	C 459	13	25.0	1074	14	US-10-156-761-2582	Sequence 2582, Ap
C 387	13	25.0	957	10	US-09-938-842A-2529	Sequence 2529, Ap	C 460	13	25.0	1084	12	US-10-027-632-261721	Sequence 261721, Ap
C 388	13	25.0	966	9	US-09-864-761-17222	Sequence 17222, A	C 461	13	25.0	1084	13	US-10-027-632-261721	Sequence 261721, Ap
C 389	13	25.0	967	12	US-10-027-632-30606	Sequence 30606, A	C 462	13	25.0	1086	9	US-09-864-761-17120	Sequence 17120, A
C 390	13	25.0	967	13	US-10-027-632-30606	Sequence 30606, A	C 463	13	25.0	1086	9	US-09-864-761-17617	Sequence 17617, A
C 391	13	25.0	967	14	US-10-106-698-521	Sequence 521, App	C 464	13	25.0	1100	10	US-09-974-300-246	Sequence 246, App
C 392	13	25.0	972	11	US-09-934-455-497	Sequence 455, App	C 465	13	25.0	1102	9	US-09-925-301-477	Sequence 477, App
C 393	13	25.0	987	14	US-10-146-772-299	Sequence 299, App	C 466	13	25.0	1102	14	US-10-106-698-342	Sequence 342, App
C 394	13	25.0	988	11	US-09-946-374-305	Sequence 305, App	C 467	13	25.0	1127	9	US-09-770-445-85	Sequence 85, Appl1
C 395	13	25.0	988	12	US-10-015-387A-305	Sequence 305, App	C 468	13	25.0	1127	10	US-09-887-576-359	Sequence 359, App
C 396	13	25.0	989	12	US-10-006-130A-305	Sequence 305, App	C 469	13	25.0	1151	10	US-09-822-830A-429	Sequence 429, App
C 397	13	25.0	989	12	US-10-006-172A-305	Sequence 305, App	C 470	13	25.0	1155	12	US-10-287-174-213	Sequence 213, App
C 398	13	25.0	989	12	US-10-015-392A-305	Sequence 305, App	C 471	13	25.0	1155	12	US-10-259-165-521	Sequence 521, App
C 399	13	25.0	989	12	US-10-017-253A-305	Sequence 305, App	C 472	13	25.0	1214	12	US-10-027-632-30343	Sequence 30343, App
C 400	13	25.0	989	12	US-10-017-306A-305	Sequence 305, App	C 473	13	25.0	1214	12	US-10-027-632-30344	Sequence 30344, A
C 401	13	25.0	989	12	US-10-210-951-21	Sequence 21, Appl1	C 474	13	25.0	1214	13	US-10-027-632-30344	Sequence 30344, A
C 402	13	25.0	989	12	US-10-211-884-21	Sequence 21, Appl1	C 475	13	25.0	1227	10	US-10-027-632-30344	Sequence 30344, A
C 403	13	25.0	989	12	US-10-012-064A-305	Sequence 305, App	C 476	13	25.0	1227	13	US-10-027-632-30344	Sequence 30344, A
C 404	13	25.0	989	12	US-10-017-867A-305	Sequence 305, App	C 477	13	25.0	1227	14	US-10-156-761-6201	Sequence 6201, Ap
C 405	13	25.0	989	12	US-10-012-101B-305	Sequence 305, App	C 478	13	25.0	1263	14	US-10-156-761-1370	Sequence 1370, Ap
C 406	13	25.0	989	12	US-10-012-137A-305	Sequence 305, App	C 479	13	25.0	1273	14	US-10-156-761-1370	Sequence 1370, Ap
C 407	13	25.0	989	12	US-10-012-752A-305	Sequence 305, App	C 480	13	25.0	1275	13	US-10-027-632-208120	Sequence 208120, Sequence 208120, Sequence 208120,
C 408	13	25.0	989	12	US-10-012-754A-305	Sequence 305, App	C 481	13	25.0	1275	13	US-10-027-632-208120	Sequence 4781, Ap
C 409	13	25.0	989	12	US-10-013-909A-305	Sequence 305, App	C 482	13	25.0	1275	14	US-10-156-761-19781	Sequence 1940, Ap
C 410	13	25.0	989	12	US-10-013-910A-305	Sequence 305, App	C 483	13	25.0	1280	11	US-09-988-067B-31	Sequence 31, Appl1
C 411	13	25.0	989	12	US-10-013-911A-305	Sequence 305, App	C 484	13	25.0	1281	10	US-09-881-752A-73	Sequence 73, Appl1
C 412	13	25.0	989	12	US-10-013-912A-305	Sequence 305, App	C 485	13	25.0	1281	10	US-10-156-761-7414	Sequence 7414, Ap
C 413	13	25.0	989	12	US-10-015-610A-305	Sequence 305, App	C 486	13	25.0	1287	14	US-09-938-864-378	Sequence 378, App
C 414	13	25.0	989	12	US-10-015-653A-305	Sequence 305, App	C 487	13	25.0	1291	11	US-09-938-864-378	Sequence 378, App
C 415	13	25.0	989	12	US-10-015-671A-305	Sequence 305, App	C 488	13	25.0	1291	12	US-10-195-835-378	Sequence 378, App
C 416	13	25.0	989	12	US-10-012-237A-305	Sequence 305, App	C 489	13	25.0	1291	14	US-10-125-635A-378	Sequence 378, App
C 417	13	25.0	989	12	US-10-013-906A-305	Sequence 305, App	C 490	13	25.0	1292	11	US-09-938-864-377	Sequence 377, App
C 418	13	25.0	989	12	US-10-015-480A-305	Sequence 305, App	C 491	13	25.0	1292	12	US-10-195-835-377	Sequence 377, App
C 419	13	25.0	989	12	US-10-015-488A-305	Sequence 305, App	C 492	13	25.0	1292	14	US-10-125-635A-377	Sequence 377, App
C 420	13	25.0	989	12	US-10-015-715A-305	Sequence 305, App	C 493	13	25.0	1302	9	US-09-879-957-17	Sequence 377, App
C 421	13	25.0	989	12	US-10-012-753A-305	Sequence 305, App	C 494	13	25.0	1302	9	US-09-879-957-17	Sequence 377, App
C 422	13	25.0	989	12	US-10-015-385A-305	Sequence 305, App	C 495	13	25.0	1311	12	US-10-027-632-123857	Sequence 123857, Sequence 123857, Sequence 123857,
C 423	13	25.0	989	12	US-10-007-236A-305	Sequence 305, App	C 496	13	25.0	1311	12	US-10-027-632-123857	Sequence 123857, Sequence 123857, Sequence 123857,
C 424	13	25.0	989	12	US-10-015-389A-305	Sequence 305, App	C 497	13	25.0	1311	12	US-10-027-632-123857	Sequence 123857, Sequence 123857, Sequence 123857,
C 425	13	25.0	989	12	US-10-015-393A-305	Sequence 305, App	C 498	13	25.0	1311	12	US-10-027-632-123857	Sequence 123857, Sequence 123857, Sequence 123857,
C 426	13	25.0	989	12	US-10-015-394A-305	Sequence 305, App	C 499	13	25.0	1311	12	US-10-027-632-123857	Sequence 123857, Sequence 123857, Sequence 123857,
C 427	13	25.0	989	12	US-10-015-519A-305	Sequence 305, App	C 500	13	25.0	1311	13	US-10-027-632-123857	Sequence 123857, Sequence 123857, Sequence 123857,
C 428	13	25.0	989	14	US-10-006-856A-305	Sequence 305, App	C 501	13	25.0	1311	13	US-10-027-632-123857	Sequence 123857, Sequence 123857, Sequence 123857,
C 429	13	25.0	989	14	US-10-006-856A-305	Sequence 305, App	C 502	13	25.0	1311	14	US-10-081-872-13	Sequence 13, Appl1
C 430	13	25.0	989	14	US-10-015-869A-305	Sequence 305, App	C 503	13	25.0	1311	14	US-10-081-872-13	Sequence 13, Appl1
C 431	13	25.0	989	14	US-10-012-121A-305	Sequence 305, App	C 504	13	25.0	1311	14	US-10-081-872-13	Sequence 13, Appl1
C 432	13	25.0	989	14	US-10-012-121A-305	Sequence 305, App	C 505	13	25.0	1311	14	US-10-081-872-13	Sequence 13, Appl1
C 433	13	25.0	989	14	US-10-006-117A-305	Sequence 305, App	C 506	13	25.0	1311	14	US-10-081-872-13	Sequence 13, Appl1
C 434	13	25.0	989	14	US-10-006-117A-305	Sequence 305, App	C 507	13	25.0	1311	14	US-10-081-872-47	Sequence 47, Appl1
C 435	13	25.0	989	14	US-10-017-527A-305	Sequence 305, App	C 508	13	25.0	1311	14	US-10-081-872-51	Sequence 51, Appl1
C 436	13	25.0	989	14	US-10-013-913A-305	Sequence 305, App	C 509	13	25.0	1311	14	US-10-081-872-53	Sequence 53, Appl1
C 437	13	25.0	989	14	US-10-007-194A-305	Sequence 305, App	C 510	13	25.0	1311	14	US-10-081-872-57	Sequence 57, Appl1
C 438	13	25.0	989	14	US-10-013-430A-305	Sequence 305, App	C 511	13	25.0	1311	14	US-10-081-872-63	Sequence 63, Appl1
C 439	13	25.0	989	14	US-10-011-671A-305	Sequence 305, App	C 512	13	25.0	1317	12	US-10-252-157-174	Sequence 174, App
C 440	13	25.0	989	14	US-10-012-755A-305	Sequence 305, App	C 513	13	25.0	1317	12	US-10-252-157-174	Sequence 174, App
C 441	13	25.0	989	14	US-10-015-386A-305	Sequence 305, App	C 514	13	25.0	1338	14	US-10-156-761-5115	Sequence 5115, Ap
C 442	13	25.0	989	14	US-10-011-692A-305	Sequence 305, App	C 515	13	25.0	1338	14	US-10-156-761-5115	Sequence 5115, Ap
C 443	13	25.0	989	14	US-10-006-768A-305	Sequence 305, App	C 516	13	25.0	1338	14	US-10-156-761-5115	Sequence 5115, Ap
C 444	13	25.0	989	14	US-10-017-610A-305	Sequence 305, App	C 517	13	25.0	1338	14	US-10-156-761-5115	Sequence 5115, Ap
C 445	13	25.0	989	14	US-10-006-063A-305	Sequence 305, App	C 518	13	25.0	1338	14	US-10-156-761-5115	Sequence 5115, Ap
C 446	13	25.0	989	14	US-10-020-063A-305	Sequence 305, App	C 519	13	25.0	1338	14	US-10-156-761-5115	Sequence 5115, Ap
C 447	13	25.0	989	14	US-10-015-391A-305	Sequence 305, App	C 520	13	25.0	1338	14	US-10-156-761-5115	Sequence 5115, Ap
C 448	13	25.0	989	14	US-10-017-407A-305	Sequence 305, App	C 521	13	25.0	1338	14	US-10-156-761-5115	Sequence 5115, Ap
C 449	13	25.0	989	15	US-10-006-041A-305	Sequence 305, App	C 522	13	25.0	1338	14	US-10-081-872A-81	Sequence 81, Appl1
C 450	13	25.0	989	15	US-10-011-813A-305	Sequence 305, App	C 523	13	25.0	1338	14	US-10-081-872A-81	Sequence 81, Appl1
C 451	13	25.0	989	15	US-10-015-822A-305	Sequence 305, App	C 524	13	25.0	1338	14	US-10-081-872A-81	Sequence 81, Appl1
C 452	13	25.0	989	15	US-10-015-822A-305	Sequence 305, App	C 525	13	25.0	1338	14	US-10-081-872A-81	Sequence 81, Appl1
C 453	13	25.0	1024	12	US-10-210-951-43	Sequence 43, Appl1	C 526	13	25.0	1338	14	US-10-081-872A-81	Sequence 81, Appl1
C 454	13	25.0	1029	14	US-10-156-761-5574	Sequence 5574, Ap	C 527	13	25.0	1338	14	US-10-142-231-53	Sequence 53, Appl1

C 528	13	25.0	1547	10	US-09-978-295A-465	Sequence 465, App	C 601	13	25.0	1547	12	US-10-141-700-417	Sequence 417, App
C 528	13	25.0	1547	10	US-09-978-697-465	Sequence 465, App	C 602	13	25.0	1547	12	US-10-141-705-417	Sequence 417, App
C 530	13	25.0	1547	10	US-09-978-192A-465	Sequence 465, App	C 603	13	25.0	1547	12	US-10-141-753-417	Sequence 417, App
C 531	13	25.0	1547	10	US-09-999-832A-465	Sequence 465, App	C 604	13	25.0	1547	12	US-10-141-758-417	Sequence 417, App
C 532	13	25.0	1547	11	US-09-978-189-465	Sequence 465, App	C 605	13	25.0	1547	12	US-10-142-118-417	Sequence 417, App
C 533	13	25.0	1547	11	US-09-978-608A-465	Sequence 465, App	C 606	13	25.0	1547	12	US-10-142-420-417	Sequence 417, App
C 534	13	25.0	1547	11	US-09-978-508A-465	Sequence 465, App	C 607	13	25.0	1547	12	US-10-142-422-417	Sequence 417, App
C 535	13	25.0	1547	11	US-09-978-191A-465	Sequence 465, App	C 608	13	25.0	1547	12	US-10-142-427-417	Sequence 417, App
C 536	13	25.0	1547	11	US-09-978-403A-465	Sequence 465, App	C 609	13	25.0	1547	12	US-10-142-760-417	Sequence 417, App
C 537	13	25.0	1547	11	US-09-978-564A-465	Sequence 465, App	C 610	13	25.0	1547	12	US-10-143-030A-465	Sequence 465, App
C 538	13	25.0	1547	11	US-09-999-833A-465	Sequence 465, App	C 611	13	25.0	1547	12	US-10-145-821-417	Sequence 417, App
C 539	13	25.0	1547	11	US-09-981-915A-465	Sequence 465, App	C 612	13	25.0	1547	12	US-10-152-831-417	Sequence 417, App
C 540	13	25.0	1547	11	US-09-978-824-465	Sequence 465, App	C 613	13	25.0	1547	12	US-10-147-940A-417	Sequence 417, App
C 541	13	25.0	1547	11	US-09-918-858A-465	Sequence 465, App	C 614	13	25.0	1547	12	US-10-142-424-417	Sequence 417, App
C 542	13	25.0	1547	11	US-09-978-423A-465	Sequence 465, App	C 615	13	25.0	1547	12	US-10-142-761-417	Sequence 417, App
C 543	13	25.0	1547	11	US-09-978-193A-465	Sequence 465, App	C 616	13	25.0	1547	12	US-10-142-763-417	Sequence 417, App
C 544	13	25.0	1547	11	US-09-999-830A-465	Sequence 465, App	C 617	13	25.0	1547	12	US-10-142-765-417	Sequence 417, App
C 545	13	25.0	1547	11	US-09-978-757A-465	Sequence 465, App	C 618	13	25.0	1547	12	US-10-142-887-417	Sequence 417, App
C 546	13	25.0	1547	11	US-09-978-187B-465	Sequence 465, App	C 619	13	25.0	1547	12	US-10-142-888-417	Sequence 417, App
C 547	13	25.0	1547	11	US-09-978-643A-465	Sequence 465, App	C 620	13	25.0	1547	12	US-10-143-034-417	Sequence 417, App
C 548	13	25.0	1547	12	US-09-978-375A-465	Sequence 465, App	C 621	13	25.0	1547	12	US-10-143-116-417	Sequence 417, App
C 549	13	25.0	1547	12	US-09-978-188A-465	Sequence 465, App	C 622	13	25.0	1547	12	US-10-143-117-417	Sequence 417, App
C 550	13	25.0	1547	12	US-09-978-298A-465	Sequence 465, App	C 623	13	25.0	1547	12	US-10-145-117-417	Sequence 417, App
C 551	13	25.0	1547	12	US-10-137-870-417	Sequence 417, App	C 624	13	25.0	1547	12	US-10-144-992-417	Sequence 417, App
C 552	13	25.0	1547	12	US-10-140-018-417	Sequence 417, App	C 625	13	25.0	1547	12	US-10-145-015-417	Sequence 417, App
C 553	13	25.0	1547	12	US-10-140-021-417	Sequence 417, App	C 626	13	25.0	1547	12	US-10-145-090-417	Sequence 417, App
C 554	13	25.0	1547	12	US-10-140-024-417	Sequence 417, App	C 627	13	25.0	1547	12	US	


```
C 966 13 25.0 1547 14 US-10-127-824A-417 Sequence 417, App
C 967 13 25.0 1547 14 US-10-127-826A-417 Sequence 417, App
C 968 13 25.0 1547 14 US-10-127-827A-417 Sequence 417, App
C 969 13 25.0 1547 14 US-10-127-828A-417 Sequence 417, App
C 970 13 25.0 1547 14 US-10-127-830A-417 Sequence 417, App
C 971 13 25.0 1547 14 US-10-127-832A-417 Sequence 417, App
C 972 13 25.0 1547 14 US-10-127-833A-417 Sequence 417, App
C 973 13 25.0 1547 14 US-10-127-834A-417 Sequence 417, App
C 974 13 25.0 1547 14 US-10-127-836A-417 Sequence 417, App
C 975 13 25.0 1547 14 US-10-127-841A-417 Sequence 417, App
C 976 13 25.0 1547 14 US-10-127-844A-417 Sequence 417, App
C 977 13 25.0 1547 14 US-10-128-687A-417 Sequence 417, App
C 978 13 25.0 1547 14 US-10-128-688A-417 Sequence 417, App
C 979 13 25.0 1547 14 US-10-128-689A-417 Sequence 417, App
C 980 13 25.0 1547 14 US-10-128-694A-417 Sequence 417, App
C 981 13 25.0 1547 14 US-10-131-825A-417 Sequence 417, App
C 982 13 25.0 1547 14 US-10-131-828A-417 Sequence 417, App
C 983 13 25.0 1547 14 US-10-131-815A-417 Sequence 417, App
C 984 13 25.0 1547 14 US-10-131-817A-417 Sequence 417, App
C 985 13 25.0 1547 14 US-10-131-821A-417 Sequence 417, App
C 986 13 25.0 1547 14 US-10-131-822A-417 Sequence 417, App
C 987 13 25.0 1547 14 US-10-131-828A-417 Sequence 417, App
C 988 13 25.0 1547 14 US-10-131-835A-417 Sequence 417, App
C 989 13 25.0 1547 14 US-10-137-864A-417 Sequence 417, App
C 990 13 25.0 1547 14 US-10-137-869A-417 Sequence 417, App
C 991 13 25.0 1547 14 US-10-147-523-417 Sequence 417, App
C 992 13 25.0 1547 14 US-10-158-785-417 Sequence 417, App
C 993 13 25.0 1547 14 US-10-121-051-417 Sequence 417, App
C 994 13 25.0 1547 14 US-10-121-042-417 Sequence 417, App
C 995 13 25.0 1547 14 US-10-123-912-417 Sequence 417, App
C 996 13 25.0 1547 14 US-10-166-709A-465 Sequence 465, App
C 997 13 25.0 1547 14 US-10-192-007-417 Sequence 417, App
C 998 13 25.0 1547 14 US-10-194-359-417 Sequence 417, App
C 999 13 25.0 1547 14 US-10-127-847A-417 Sequence 417, App
C1000 13 25.0 1547 14 US-10-175-590-417 Sequence 417, App
```

ALIGNMENTS

```
RESULT 1
US-10-286-606-11/c
; Sequence 11, Application US/10286606
; Publication No. US20030180321A1
; GENERAL INFORMATION:
; APPLICANT: Bertozzi, Carolyn
; APPLICANT: Williams, Spencer J.
; APPLICANT: Mougous, Joseph
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins
; TITLE OF INVENTION: and Methods of Use Thereof
; FILE REFERENCE: BERK-012CIP
; CURRENT APPLICATION NUMBER: US/10/286,606
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: 60/285,394
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/345,953
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 10/126,279
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Mycobacterium avium
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1392)
US-10-286-606-11
```

```
Query Match 32.7%; Score 17; DB 12; Length 1392;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 29 TCCTCGCTGCCGCGCTC 45
|||||
Db 290 TCCTCGCTGCCGCGCTC 274
```

```
RESULT 2
US-10-126-279-11/c
; Sequence 11, Application US/10126279
; Publication No. US20030104001A1
; GENERAL INFORMATION:
; APPLICANT: Bertozzi, Carolyn
; APPLICANT: Williams, Spencer J.
; APPLICANT: Mougous, Joseph
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins
; TITLE OF INVENTION: and Methods of Use Thereof
; FILE REFERENCE: BERK-012
; CURRENT APPLICATION NUMBER: US/10/126,279
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/285,394
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/345,953
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Mycobacterium avium
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1392)
US-10-126-279-11
```

```
Query Match 32.7%; Score 17; DB 14; Length 1392;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 29 TCCTCGCTGCCGCGCTC 45
|||||
Db 290 TCCTCGCTGCCGCGCTC 274
```

```
RESULT 3
US-09-918-995-25977
; Sequence 25977, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25977
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(574)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-25977
```

```
Query Match 30.8%; Score 16; DB 11; Length 574;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 19 CTCATCGCTCTCTCTCG 34
```

```
Db          194 CTCATCGCTCTCGTCG 209
|||||
RESULT 4
US-09-897-214-7
; Sequence 7, Application US/09897214
; Patent No. US20020076779A1
; GENERAL INFORMATION:
; APPLICANT: Thayer, Edward C.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: Leucine-Rich Repeat Proteins, Zlrr7,
; FILE REFERENCE: 01-27
; CURRENT APPLICATION NUMBER: US/09/897,214
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1653
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1653)
US-09-897-214-7

Query Match          30.8%; Score 16; DB 9; Length 1653;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          33 CGCTGCGGCGCTCGCC 48
|||||
Db          1496 CGCTGCGGCGCTCGCC 1511

RESULT 5
US-09-768-826-21
; Sequence 21, Application US/09768826
; Patent No. US20020012966A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; FILE REFERENCE: PFS12P1
; TITLE OF INVENTION: 18 human secreted proteins
; CURRENT APPLICATION NUMBER: US/09/768,826
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US00/22350
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/148,759
; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-768-826-21

Query Match          30.8%; Score 16; DB 9; Length 1872;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          33 CGCTGCGGCGCTCGCC 48
|||||
Db          1153 CGCTGCGGCGCTCGCC 1168

RESULT 6
US-10-093-463-189
; Sequence 189, Application US/10093463
; Publication No. US20030208039A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glenda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangolli, Beha
; APPLICANT: Vernet, Corine
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malysankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Paturnajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: No. US20030208039A1e1 Antibodies that Bind to Antigenic Polypepi
; FILE REFERENCE: 21402-290A (Cura 590AT)
; CURRENT APPLICATION NUMBER: US/10/093,463
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/283,675
; PRIOR FILING DATE: 2001-04-14
; PRIOR APPLICATION NUMBER: 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,101
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/275,578
; PRIOR FILING DATE: 2001-03-13
```



```
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 189
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (471)..(2169)
US-10-093-463-189

Query Match          30.8%; Score 16; DB 12; Length 2451;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33  CGCTGCCGCGCTGCC 48
Db      1966 CGCTGCCGCGCTGCC 1981

RESULT 7
US-10-093-463-191
; Sequence 191, Application US/10093463
; Publication No. US20030208039A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glenda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corine
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malyankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Paturajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: No. US20030208039A1 Antibodies that Bind to Antigenic Polypept
; FILE REFERENCE: 21402-290A (Cura 590AT)
; CURRENT APPLICATION NUMBER: US/10/093,463
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/283,675
; PRIOR FILING DATE: 2001-04-14
; PRIOR APPLICATION NUMBER: 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,101
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
```

```
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/275,578
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 191
; LENGTH: 2563
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (130)..(2464)
US-10-093-463-191

Query Match          30.8%; Score 16; DB 12; Length 2563;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33  CGCTGCCGCGCTGCC 48
Db      1679 CGCTGCCGCGCTGCC 1694

RESULT 8
US-10-246-330-1
; Sequence 1, Application US/10246330
; Publication No. US20030166030A1
; GENERAL INFORMATION:
; APPLICANT: O'Toole, George A.
; APPLICANT: Mah, Thien-Pah
; TITLE OF INVENTION: METHODS TO STUDY AND MECHANISMS OF
; FILE REFERENCE: 14537-002001
; CURRENT APPLICATION NUMBER: US/10/246,330
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/323,241
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2610
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2607)
US-10-246-330-1

Query Match          30.8%; Score 16; DB 12; Length 2610;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      25  GCTCTGCTGCTGCCG 40
```

Db 2547 GCCTCGCTCGCTCGC 2562

RESULT 9

US-10-093-463-187
Sequence 187, Application US/10093463
Publication No. US20030208039A1

GENERAL INFORMATION:

APPLICANT: Padigaru, Muralidhara

APPLICANT: Shenoy, Suresh

APPLICANT: Khekda, Ramesh

APPLICANT: Gusev, Vladimir

APPLICANT: Pochart, Pascal

APPLICANT: Zhong, Mei

APPLICANT: Rastelli, Luca

APPLICANT: Wezes, Peter

APPLICANT: Smithson, Glenda

APPLICANT: Guo, Xiaojia

APPLICANT: Gerlach, Valerie

APPLICANT: Casman, Stacie

APPLICANT: Boldog, Ferenc

APPLICANT: Li, Li

APPLICANT: Zerhusen, Bryan

APPLICANT: Tchernev, Velizar

APPLICANT: Gangolli, Esha

APPLICANT: Vernet, Corine

APPLICANT: Pena, Carol

APPLICANT: Burgess, Catherine

APPLICANT: Liu, Xiaohong

APPLICANT: Spytek, Kimberly

APPLICANT: Gorman, Linda

APPLICANT: Spaderna, Steven

APPLICANT: Voss, Edward

APPLICANT: Malyankar, Uriel

APPLICANT: Anderson, David

APPLICANT: Patlurajan, Meera

APPLICANT: Miller, Charles

APPLICANT: Taupier, Raymond J., Jr.

TITLE OF INVENTION: No. US20030208039A1 Antibodies that Bind to Antigenic Polypeptide

FILE REFERENCE: 21402-290A (Cura 590A1)

CURRENT FILING DATE: 2002-06-24

PRIOR APPLICATION NUMBER: 60/283,675

PRIOR FILING DATE: 2001-04-14

PRIOR APPLICATION NUMBER: 60/338,092

PRIOR FILING DATE: 2001-12-03

PRIOR APPLICATION NUMBER: 60/274,281

PRIOR FILING DATE: 2001-03-08

PRIOR APPLICATION NUMBER: 60/274,101

PRIOR FILING DATE: 2001-03-08

PRIOR APPLICATION NUMBER: 60/325,681

PRIOR FILING DATE: 2001-09-27

PRIOR APPLICATION NUMBER: 60/304,354

PRIOR FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: 60/279,995

PRIOR FILING DATE: 2001-03-30

PRIOR APPLICATION NUMBER: 60/294,899

PRIOR FILING DATE: 2001-05-31

PRIOR APPLICATION NUMBER: 60/287,424

PRIOR FILING DATE: 2001-04-30

PRIOR APPLICATION NUMBER: 60/299,027

PRIOR FILING DATE: 2001-06-18

PRIOR APPLICATION NUMBER: 60/309,198

PRIOR FILING DATE: 2001-07-31

PRIOR APPLICATION NUMBER: 60/281,194

PRIOR FILING DATE: 2001-04-04

PRIOR APPLICATION NUMBER: 60/274,194

PRIOR FILING DATE: 2001-03-08

PRIOR APPLICATION NUMBER: 60/274,849

PRIOR FILING DATE: 2001-03-09

PRIOR APPLICATION NUMBER: 60/330,380

PRIOR FILING DATE: 2001-10-18

PRIOR APPLICATION NUMBER: 60/275,235

PRIOR FILING DATE: 2001-03-12

PRIOR APPLICATION NUMBER: 60/288,342

PRIOR FILING DATE: 2001-05-03

PRIOR APPLICATION NUMBER: 60/275,578

PRIOR FILING DATE: 2001-03-13

NUMBER OF SEQ ID NOS: 370

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 187

LENGTH: 3201

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (471)..(2412)

US-10-093-463-187

Query Match 30.8%; Score 16; DB 12; Length 3201;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CGCTGCCGCGCTCGCC 48

Db 1966 CGCTGCCGCGCTCGCC 1981

RESULT 10

US-09-925-300-638/c

Sequence 638, Application US/09925300

Patent No. US20020151681A1

GENERAL INFORMATION:

APPLICANT: Craig Rosen,

APPLICANT: Steve Ruben,

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA101

CURRENT APPLICATION NUMBER: US/09/925,300

PRIOR FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05988

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 1890

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 638

LENGTH: 3947

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (625)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (3738)

OTHER INFORMATION: n equals a,t,g, or c

US-09-925-300-638

Query Match 30.8%; Score 16; DB 10; Length 3947;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 CTCATGCTCTCTCTCG 34

Db 72 CTCATGCTCTCTCTCG 57

RESULT 11

US-10-106-698-1931/c

Sequence 1931, Application US/10106698

Publication No. US20030109690A1

GENERAL INFORMATION:

APPLICANT: Ruben et al.

TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide

```
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 1931
; LENGTH: 3953
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (618)..(618)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (626)..(626)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-1931
```

```
Query Match          30.8%; Score 16; DB 14; Length 3953;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      19 CTCATCGCTCTCGTCG 34
          |||||
Db       72 CTCATCGCTCTCGTCG 57
```

```
RESULT 12
US-09-294-093B-4207
; Sequence 4207, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 4207
; LENGTH: 214
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700353417H1
; NAME/KEY: unsure
; LOCATION: 10, 37, 204
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-4207
```

```
Query Match          28.8%; Score 15; DB 9; Length 214;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ATCAGTTCTCTCAC 15
          |||||
Db       182 ATCAGTTCTCTCAC 196
```

```
RESULT 13
US-09-878-574-15763/c
; Sequence 15763, Application US/09878574
```

```
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plante
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 15763
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701070483H1
US-09-878-574-15763
```

```
Query Match          28.8%; Score 15; DB 10; Length 264;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      26 CTCCTCGCTGCGCG 40
          |||||
Db       111 CTCCTCGCTGCGCG 97
```

```
RESULT 14
US-10-156-761-711/c
; Sequence 711, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 711
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(294)
US-10-156-761-711
```

```
Query Match          28.8%; Score 15; DB 14; Length 294;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      29 TCGTGGCTGCGCGCC 43
          |||||
Db       143 TCGTGGCTGCGCGCC 129
```

```
RESULT 15
US-09-878-574-1259/c
; Sequence 1259, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
```

APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO: 1259
LENGTH: 336
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: LIB3028-040-Q1-B1-A1
US-09-878-574-1259

Query Match 28.8%; Score 15; DB 10; Length 336;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 CTCACGCTCTCAT 23
DB 67 CTCACGCTCTCAT 53

RESULT 16
US-09-764-870-129
Sequence 129, Application US/09764870
Patent No. US20020042386A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT214
CURRENT APPLICATION NUMBER: US/09/764,870
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 646
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO: 129
LENGTH: 357
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-870-129

Query Match 28.8%; Score 15; DB 9; Length 357;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACTTCTCTCAC 15
DB 102 ATGCACTTCTCTCAC 116

RESULT 17
US-10-125-540-129
Sequence 129, Application US/10125540
Publication No. US20030059875A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT214CT
CURRENT APPLICATION NUMBER: US/10/125,540
CURRENT FILING DATE: 2002-04-19
Prior application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 646
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO: 129
LENGTH: 357
TYPE: DNA
ORGANISM: Homo sapiens
US-10-125-540-129

Query Match 28.8%; Score 15; DB 14; Length 357;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACTTCTCTCAC 15
DB 102 ATGCACTTCTCTCAC 116

RESULT 18
US-10-084-843-264/C
Sequence 264, Application US/10084843
Publication No. US20030143243A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
Skeiky, Yasir A.W.
Dillon, David C.
Campos-Neto, Antonio
Houghton, Raymond
Vedvick, Thomas S.
Twardzik, Daniel R.
Lodes, Michael J.
Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 355
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/084,843
FILING DATE: 25-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 264:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 264:
US-10-084-843-264

Query Match 28.8%; Score 15; DB 12; Length 423;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 GGTGCTGCGGCGCT 44
DB 26 GTGCTGCGGCGCT 12

RESULT 19

```
US-10-193-002-259/c
; Sequence 259, Application US/10193002
; Publication No. US20030135026A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
;           Skeiky, Yasir A.W.
;           Dillon, Davin C.
;           Campos-Neto, Antonia
;           Houghton, Raymond
;           Vedvick, Thomas S.
;           Twardzik, Daniel R.
;           Lodes, Michael J.
;           Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                   TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SEED and BERRY LLP
;   STREET: 6300 Columbia Center, 701 Fifth Avenue
;   CITY: Seattle
;   STATE: Washington
;   COUNTRY: USA
;   ZIP: 98104-7092
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/193,002
;   FILING DATE: 10-Jul-2002
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/09/072,596
;     FILING DATE: 05-MAY-1998
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Maki, David J.
;     REGISTRATION NUMBER: 31,392
;     REFERENCE/DOCKET NUMBER: 210121.417C9
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (206) 622-4900
;     TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 259:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 423 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     SEQUENCE DESCRIPTION: SEQ ID NO: 259:
US-10-193-002-259
Query Match      28.8%; Score 15; DB 12; Length 423;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy      30 CGTCGCTGCCGCGCT 44
Db      26 CGTCGCTGCCGCGCT 12
RESULT 20
US-10-259-678-60
; Sequence 60, Application US/10259678
; Publication No. US20030198974A1
; GENERAL INFORMATION:
; APPLICANT: Cole, Stewart
;           Buchrieser-Brosch, Roland
;           Gordon, Stephen
;           Billault, Alain
; TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
;                   THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
; TITLE OF INVENTION: LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA
```

```
; FILE REFERENCE: 3495-0169
; CURRENT APPLICATION NUMBER: US/10/259,678
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/060,756
; PRIOR FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 60
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (various positions within the sequence)
;   OTHER INFORMATION: applicants are uncertain of bases designated as "n"
US-10-259-678-60
Query Match      28.8%; Score 15; DB 12; Length 448;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy      7 TTCTCAGCGCTCTC 21
Db      352 TTCTCAGCGCTCTC 366
RESULT 21
US-09-864-761-6099
; Sequence 6099, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
;           Rank, David R.
;           Hanzel, David K.
;           Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;                   FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
```

```

; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6099
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004923.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 16
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 14
US-09-864-761-6099
```

```

Query Match      28.8%: Score 15; DB 9; Length 481;
Best Local Similarity 100.0%: Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      15  CGCTCTCATGCTCT 29
      |||||
DB      152 CGCTCTCATGCTCT 166
```

```

RESULT 22
US-10-027-632-115584
; Sequence 115584, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115584
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115584
```

```

Query Match      28.8%: Score 15; DB 12; Length 603;
Best Local Similarity 100.0%: Pred. No. 69;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      25  GCTCTGTCGCTGCC 39
      |||||
DB      48  GCTCTGTCGCTGCC 62
```

RESULT 23

```

US-10-027-632-115584
; Sequence 115584, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115584
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115584
```

```

Query Match      28.8%: Score 15; DB 13; Length 603;
Best Local Similarity 100.0%: Pred. No. 69;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      25  GCTCTGTCGCTGCC 39
      |||||
DB      48  GCTCTGTCGCTGCC 62
```

```

RESULT 24
US-10-156-761-974/c
; Sequence 974, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 974
; LENGTH: 1113
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1113)
US-10-156-761-974
```

```

Query Match      28.8%: Score 15; DB 14; Length 1113;
Best Local Similarity 100.0%: Pred. No. 63;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 29 TCGTCGCTGCCGCC 43
|||||
Db 845 TCGTCGCTGCCGCC 841

RESULT 25

US-10-259-165-279
; Sequence 279, Application US/10259165
; Publication No. US2003013588A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Tong
; APPLICANT: Wang, Xun
; APPLICANT: Chang, Hur-song
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyuki
; APPLICANT: Kieps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Rieke, Darrell
; TITLE OF INVENTION: GENES THAT ARE MODULATED BY POSTTRANSCRIPTIONAL GENE SILENCING
; FILE REFERENCE: 70030-NP
; CURRENT APPLICATION NUMBER: US/10/259,165
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/368,327
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 279
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-259-165-279

Query Match 28.8%; Score 15; DB 12; Length 1191;
Best Local Similarity 100.0%; Pred. No. 63;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 TCGCTGCCGCCCTCG 46
|||||
Db 88 TCGCTGCCGCCCTCG 102

RESULT 26

US-10-037-270-168/c
; Sequence 168, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Aseundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_fl_genes Version 1.0
; SEQ ID NO 168
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)..(548)
US-10-037-270-168

Query Match 28.8%; Score 15; DB 14; Length 1197;
Best Local Similarity 100.0%; Pred. No. 63;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 CTCATCGCTCTGTC 33
|||||
Db 100 CTCATCGCTCTGTC 86

RESULT 27

US-10-156-761-780/c
; Sequence 780, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 780
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Streptomyces avermectilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1389)
US-10-156-761-780

Query Match 28.8%; Score 15; DB 14; Length 1389;
Best Local Similarity 100.0%; Pred. No. 61;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 CGTCGCTGCCGCCCT 44
|||||
Db 19 CGTCGCTGCCGCCCT 5

RESULT 28

US-10-177-293-416/c
; Sequence 416, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James

```

; APPLICANT: Glatz, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobaagi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 416
; LENGTH: 1404
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-416

Query Match      28.8%; Score 15; DB 14; Length 1404;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 CTCATCGCTCTGCTC 33
Db      122 CTCATCGCTCTGCTC 108

RESULT 29
US-10-156-761-4461/c
; Sequence 4461, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OKURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 4461
```

```

; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1479)
US-10-156-761-4461

Query Match      28.8%; Score 15; DB 14; Length 1479;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      27 TCTCGTGGTGGCCG 41
Db      1258 TCTCGTGGTGGCCG 1244

RESULT 30
US-10-128-714-7070/c
; Sequence 7070, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Erosnkin, Alexey M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: Methods of Use
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7070
; LENGTH: 1554
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-7070

Query Match      28.8%; Score 15; DB 14; Length 1554;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      24 CGCTCGTGGCTGCG 38
Db      1384 CGCTCGTGGCTGCG 1370

RESULT 31
US-10-128-714-6070/c
; Sequence 6070, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Erosnkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: Methods of Use
```



```
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128, 714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285, 697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287, 066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295, 890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303, 899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316, 362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6070
LENGTH: 1778
TYPE: DNA
ORGANISM: Aspergillus fumigatus
US-10-128-714-6070
```

```
Query Match      28.8%; Score 15; DB 14; Length 1778;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      24 CGCTCTGCTGCTGC 38
      |||
Db      1608 CGCTCTGCTGCTGC 1594
```

```
RESULT 32
US-09-801-368-235/c
Sequence 235, Application US/09801368
Patent No. US20020128250A1
GENERAL INFORMATION:
APPLICANT: Busby, Robert
APPLICANT: Cali, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Solie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801, 368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487, 558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160, 587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
SEQ ID NO 235
LENGTH: 1932
TYPE: DNA
ORGANISM: Penicillium chrysogenum
US-09-801-368-235
```

```
Query Match      28.8%; Score 15; DB 10; Length 1932;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      18 TCTCATGCTGCTGCT 32
      |||
Db      1024 TCTCATGCTGCTGCT 1010
```

```
RESULT 33
US-09-938-842A-2978
Sequence 2978, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Jun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938, 842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227, 866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264, 647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300, 111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2978
LENGTH: 2000
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2978
```

```
Query Match      28.8%; Score 15; DB 10; Length 2000;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      8 TCTCTCAGCTCTCA 22
      |||
Db      1490 TCTCTCAGCTCTCA 1504
```

```
RESULT 34
US-10-156-761-158
Sequence 158, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, WASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156, 761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 158
LENGTH: 2412
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2412)
US-10-156-761-158
```

```
Query Match      28.8%; Score 15; DB 14; Length 2412;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      28 CTGTCGCTGCGGC 42
      |||
Db      1471 CTGTCGCTGCGGC 1485
```

```
RESULT 35
US-09-815-242-4138
; Sequence 4138, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4138
; LENGTH: 2454
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-4138

Query Match      28.8%; Score 15; DB 9; Length 2454;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      30 CGTCGCTGCGGCGCT 44
Db      12 CGTCGCTGCGGCGCT 26

RESULT 36
US-09-374-046A-125
; Sequence 125, Application US/09374046A
; Publication No. US20030096951A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steindinger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fecthel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6075-83A
; CURRENT APPLICATION NUMBER: US/09/374,046A
```

```
; CURRENT FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 125
; LENGTH: 2672
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-374-046A-125

Query Match      28.8%; Score 15; DB 11; Length 2672;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      17 CTTCTCATCGCTCTCG 31
Db      153 CTTCTCATCGCTCTCG 167

RESULT 37
US-09-954-456-518/C
; Sequence 518, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Car
; TITLE OF INVENTION: Set8
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 518
; LENGTH: 2790
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-518

Query Match      28.8%; Score 15; DB 10; Length 2790;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      15 CGCTTCATCGCTCT 29
Db      984 CGCTTCATCGCTCT 970

RESULT 38
US-09-967-768A-182/C
; Sequence 182, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signat
```

```

; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 182
; LENGTH: 2790
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-967-768A-182

Query Match      28.8%; Score 15; DB 10; Length 2790;
Best Local Similarity 100.0%; Pred.No. 56;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      15  CGCTCTCATCGCTCT 29
Db      984 CGCTCTCATCGCTCT 970

RESULT 39
US-10-128-714-70/C
; Sequence 70, Application US/10128714
; Publication NO. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshekin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 3140
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
; US-10-128-714-70

Query Match      28.8%; Score 15; DB 14; Length 3140;
Best Local Similarity 100.0%; Pred.No. 55;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      24  CGCTCTCGTCTGCTGC 38
Db      2608 CGCTCTCGTCTGCTGC 2594

RESULT 40
US-09-822-846-197
; Sequence 197, Application US/09822846

```

```

; Publication No. US20030027139A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fectel, Kim
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakari
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6400
; CURRENT APPLICATION NUMBER: US/09/822,846
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,605
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 629
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 197
; LENGTH: 3323
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-822-846-197

Query Match      28.8%; Score 15; DB 11; Length 3323;
Best Local Similarity 100.0%; Pred.No. 54;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      17  CTCTCATCGCTCTCG 31
Db      153 CTCTCATCGCTCTCG 167

Search completed: November 13, 2003, 11:54:39
Job time : 246 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2003, 09:41:06 ; Search time 1948 Seconds

(without alignments)
648.785 Million cell updates/sec

Title: US-10-081-935-1

Perfect score: 52

Sequence: 1 atgcagcttcctcctcagctctc.....cgctgcgcgcgcctcgcagctg 52

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST.*
1: em_escba:*
2: em_esthum:*
3: em_escin:*
4: em_escnu:*
5: em_escov:*
6: em_escpl:*
7: em_escro:*
8: em_hic:*
9: gb_esc1:*
10: gb_esc2:*
11: gb_hic:*
12: gb_esc3:*
13: gb_esc4:*
14: gb_esc5:*
15: em_escfun:*
16: em_escrom:*
17: em_gsa_hum:*
18: em_gsa_inv:*
19: em_gsa_pin:*
20: em_gsa_vrt:*
21: em_gsa_fun:*
22: em_gsa_mam:*
23: em_gsa_mus:*
24: em_gsa_pro:*
25: em_gsa_rtd:*
26: em_gsa_phg:*
27: em_gsa_vrl:*
28: gb_gsa1:*
29: gb_gsa2:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	395	10	BF251279 EST418539
2	52	100.0	401	10	BF252856 EST420119
3	52	100.0	531	10	BF251334 EST418594
4	52	100.0	557	10	BF252723 EST419986

5	52	100.0	737	10	BF251103	BF251103 EST418360
6	23	44.2	311	13	BQ499065	BQ499065 EST088290
7	19	36.5	133	12	BP112708	BP112708 BP112708
8	19	36.5	314	9	AUI63553	AUI63553 AUI63553
9	19	36.5	460	9	AUI6766	AUI6766 AUI6766
10	19	36.5	468	9	AUI031845	AUI031845 AUI031845
11	19	36.5	585	28	AQ510461	AQ510461 nbxb00958
12	19	36.5	758	14	CB684405	CB684405 OSJNE110
13	19	36.5	767	14	CB939902	CB939902 OGCJX14
14	19	36.5	978	10	BE039817	BE039817 OCOBE11
15	19	36.5	1712	10	BG163853	BG163853 602343204
16	18	34.6	368	28	AQ845447	AQ845447 LMAJFV1.1
17	18	34.6	546	12	B1568671	B1568671 RH40331.5
18	18	34.6	558	12	B1568892	B1568892 RH22614.5
19	18	34.6	610	12	B1626323	B1626323 RH67006.5
20	18	34.6	936	28	AZ208101	AZ208101 SF_0136.B
21	18	34.6	1012	29	CNS01736	AL107532 Drosoph11
22	18	34.6	1083	10	BF307868	BF307868 601890683
23	18	34.6	1101	29	CNS00ATX	AL056234 Drosoph11
24	18	34.6	1101	29	CNS014JM	AL104246 Drosoph11
25	17	32.7	359	14	CD262900	BL123995 1033P32P
26	17	32.7	374	13	BY076213	CD262900 PMAU19XK
27	17	32.7	404	9	BF008574	BY076213 BY076213
28	17	32.7	406	10	AM465276	BF008574 1613456 A
29	17	32.7	425	9	AM056821	AM465276 BP230018B
30	17	32.7	470	9	AM056821	AM056821 ST56P10 P
31	17	32.7	490	13	BQ170403	AM376507 IL3-CT021
32	17	32.7	505	10	BE484921	BQ170403 WHE1770.E
33	17	32.7	506	13	BO821431	BE484921 171741 BA
34	17	32.7	506	13	BO821431	BO821431 1030092C0
35	17	32.7	529	9	A1988643	A1988643 gdo5ho5.Y
36	17	32.7	542	9	AV668034	AV668034 AV668034
37	17	32.7	544	9	A1542489	A1542489 S008831.5
38	17	32.7	546	9	AV597173	AV597173 AV597173
39	17	32.7	556	9	AV603191	AV603191 AV603191
40	17	32.7	570	9	AV614001	AV614001 AV614001
41	17	32.7	598	9	AV662793	AV662793 AV662793
42	17	32.7	604	28	AQ946593	AQ946593 Sheared D
43	17	32.7	616	28	BH335108	BH335108 CH230-99G
44	17	32.7	626	13	BU044311	BU044311 PP LEA001
45	17	32.7	663	10	BE776572	BE776572 MY-17-F-1
46	17	32.7	663	13	BQ162104	BQ162104 WHE0496.E
47	17	32.7	668	14	CD258920	CD258920 p1MY0175F
48	17	32.7	677	29	BZ363829	BZ363829 id91c10.G
49	17	32.7	694	14	CD491843	CD491843 WHE222.G
50	17	32.7	707	14	CB419666	CB419666 592540 MA
51	17	32.7	758	29	BZ360194	BZ360194 id91c10.b
52	17	32.7	760	14	CB653954	CB653954 OSJNEC05H
53	17	32.7	770	10	BF865131	BF865131 963057D04
54	17	32.7	771	29	CNS017FC	AL166449 Tetraodon
55	17	32.7	812	14	BY728779	BY728779 BY728779
56	17	32.7	822	29	CNS0421Q	AL171547 Tetraodon
57	17	32.7	839	10	BF701989	BF701989 602683363
58	17	32.7	960	29	CNS01EMC	AL140725 Anopheles
59	17	32.7	978	29	CNS07BHZ	AL134013 T3 end of
60	17	32.7	979	12	BM470883	BM470883 AGENCOURT
61	17	32.7	1013	29	CNS03K1A	AL156663 Tetraodon
62	17	32.7	1099	29	BZ579805	BZ579805 msh2.782.
63	17	32.7	1201	13	BX448656	BX448656 BX448656
64	17	32.7	1269	10	BE963132	BE963132 601656936
65	17	32.7	1654	29	AC0882054	AC0882054 Pan troyi
66	17	32.7	1947	10	BF303726	BF303726 602032601
67	17	32.7	113	9	AL825153	AL825153
68	16	30.8	172	9	AA665646	AA665646 ag55f03..8
69	16	30.8	188	29	DME545703	AJ345703 Drosoph11
70	16	30.8	207	29	AG215251	AG215251 Drosoph11
71	16	30.8	262	29	AJ497920	AJ497920
72	16	30.8	263	28	AQ911925	AQ911925 LMAJFV1.1
73	16	30.8	284	28	BH904772	BH904772 SALK_1050
74	16	30.8	312	29	CNS01K1L	AL144890 Anopheles
75	16	30.8	321	13	BU644056	BU644056 mgmk018XO
76	16	30.8	334	12	BMS77157	BMS77157 170006871
77	16	30.8				

C 78	16	30.8	352	12	B1023345	B1023345 CM4-MT024	C 151	16	30.8	621	9	AM395638	AM395638 sg73e02.y
C 79	16	30.8	376	13	B0754769	B0754769 Ebed01_SQ	C 152	16	30.8	632	9	AM067277	AM067277 683021D03
C 80	16	30.8	381	10	BG054345	BG054345 OV2_3_B06	C 153	16	30.8	635	14	CA730681	CA730681 w1d1c.pk0
C 81	16	30.8	381	28	A0846009	A0846009 LMA1FV1.1	C 154	16	30.8	638	14	CD0422724	CD0422724 psHB042XA
C 82	16	30.8	386	12	BM826922	BM826922 K-EST0059	C 155	16	30.8	645	28	BH379753	BH379753 AG-ND-150
C 83	16	30.8	387	10	BE400577	BE400577 AMB005.E0	C 156	16	30.8	646	14	CA014362	CA014362 HT11C07r
C 84	16	30.8	389	13	B0762332	B0762332 EBR001_SQ	C 157	16	30.8	648	14	CD462179	CD462179 SAI_39_CO
C 85	16	30.8	391	9	AV418656	AV418656 AVA18656	C 158	16	30.8	654	14	CD044693	CD044693 psHB015XH
C 86	16	30.8	396	9	AV422875	AV422875 AVA22875	C 159	16	30.8	656	14	CD424258	CD424258 SAI_4_B04
C 87	16	30.8	397	9	AW788982	AW788982 C00848-F	C 160	16	30.8	657	12	BM637161	BM637161 170006875
C 88	16	30.8	412	9	AV414607	AV414607 AVA14607	C 161	16	30.8	659	12	BJ279513	BJ279513 B079513
C 89	16	30.8	416	13	B0642727	B0642727 mgmK013xC	C 162	16	30.8	666	12	B1910227	B1910227 603069436
C 90	16	30.8	418	9	AV407733	AV407733 AVA07733	C 163	16	30.8	667	12	BM869294	BM869294 mgm004XB
C 91	16	30.8	419	9	AU091999	AU091999 AU091999	C 164	16	30.8	667	13	BQ294639	BQ294639 WHE2851.G
C 92	16	30.8	419	9	AV424304	AV424304 AVA24304	C 165	16	30.8	672	12	BM656320	BM656320 170006873
C 93	16	30.8	419	28	BH394286	BH394286 AG-ND-172	C 166	16	30.8	673	12	BM636353	BM636353 170006689
C 94	16	30.8	422	14	B1798708	B1798708 H14E06.E	C 167	16	30.8	680	9	AA941851	AA941851 LD27161.5
C 95	16	30.8	422	14	H30859	H30859 y079a10.r1	C 168	16	30.8	680	14	CD042626	CD042626 psHB041XJ
C 96	16	30.8	429	9	AV414050	AV414050 AVA14050	C 169	16	30.8	682	9	AB070961	AB070961 AB070961
C 97	16	30.8	438	12	BM868761	BM868761 mgm0001XP	C 170	16	30.8	685	14	CB028066	CB028066 TBEST-yd4
C 98	16	30.8	443	12	BQ048804	BQ048804 952024C12	C 171	16	30.8	685	14	CD043489	CD043489 psHB008XN
C 99	16	30.8	447	14	H30753	H30753 y079a06.r1	C 172	16	30.8	688	28	BH383752	BH383752 AG-ND-125
C 100	16	30.8	463	28	BH195324	BH195324 TC3-75F21	C 173	16	30.8	689	10	BG328825	BG328825 602427427
C 101	16	30.8	465	13	BQ764235	BQ764235 E8an01_SQ	C 174	16	30.8	689	14	CD426215	CD426215 SAI_19-GO
C 102	16	30.8	472	10	BF765337	BF765337 RC0-CS004	C 175	16	30.8	690	14	CD426404	CD426404 SAI_20_F0
C 103	16	30.8	478	9	AU070784	AU070784 AU070784	C 176	16	30.8	691	28	BH397190	BH397190 AG-ND-155
C 104	16	30.8	479	9	AL932614	AL932614 AL932614	C 177	16	30.8	694	13	BM578621	BM578621 WHE0307.G
C 105	16	30.8	485	28	BH370796	BH370796 AG-ND-145	C 178	16	30.8	696	29	AL945424	AL945424 Atxbl180p6
C 106	16	30.8	489	10	BE606956	BE606956 WHE0914.H	C 179	16	30.8	700	13	B0646769	B0646769 1112033D0
C 107	16	30.8	496	10	BE599468	BE599468 P11_87_H0	C 180	16	30.8	700	13	B0653829	B0653829 1112109C1
C 108	16	30.8	504	12	BM866352	BM866352 mgc0005XB	C 181	16	30.8	702	14	CD041858	CD041858 psHB037XM
C 109	16	30.8	510	14	CA601785	CA601785 w1r.pk000	C 182	16	30.8	704	28	AC058578	AC058578 LERAY43TR
C 110	16	30.8	510	28	B44411	B44411 HS-1059-A2-	C 183	16	30.8	711	28	BH580264	BH580264 BOHAE55TR
C 111	16	30.8	511	14	CA721357	CA721357 w4K9n.pk0	C 184	16	30.8	712	14	CD491648	CD491648 WHE2851.G
C 112	16	30.8	513	28	BH383883	BH383883 AG-ND-159	C 185	16	30.8	720	29	BZ2983502	BZ2983502 PUG1633TD
C 113	16	30.8	523	9	AM564386	AM564386 LG1_292.G	C 186	16	30.8	724	13	B0646992	B0646992 1112055A0
C 114	16	30.8	523	10	BG355483	BG355483 EMI_17-A0	C 187	16	30.8	724	13	CB674742	CB674742 OSJNE09P
C 115	16	30.8	532	12	BM793941	BM793941 K-EST0074	C 188	16	30.8	732	28	BH390350	BH390350 AG-ND-153
C 116	16	30.8	533	10	BG604719	BG604719 WHE0941.D	C 189	16	30.8	733	28	B1091229	B1091229 602856074
C 117	16	30.8	544	12	BM097544	BM097544 EBem04_SQ	C 190	16	30.8	738	29	CNS01MBP	CNS01MBP AG-ND-140
C 118	16	30.8	545	12	BM371104	BM371104 EBR004_SQ	C 191	16	30.8	741	28	BH370657	BH370657 AG-ND-166
C 119	16	30.8	546	14	CD039916	CD039916 psHB03IXC	C 192	16	30.8	742	10	BE273785	BE273785 601104636
C 120	16	30.8	546	14	CD040928	CD040928 psHB035XC	C 193	16	30.8	758	28	BH385126	BH385126 AG-ND-167
C 121	16	30.8	547	12	BM371056	BM371056 EBR004_SQ	C 194	16	30.8	765	28	BH388423	BH388423 AG-ND-162
C 122	16	30.8	551	10	BE403428	BE403428 WHE0429.D	C 195	16	30.8	766	28	BH386360	BH386360 AG-ND-146
C 123	16	30.8	557	12	BM632289	BM632289 170006875	C 196	16	30.8	772	10	BE645657	BE645657 EST507220
C 124	16	30.8	557	13	BQ317031	BQ317031 RC1-CT027	C 197	16	30.8	777	10	BE540337	BE540337 601061228
C 125	16	30.8	558	13	B0838378	B0838378 WHE2909-G	C 198	16	30.8	803	10	BF529190	BF529190 602041622
C 126	16	30.8	558	14	CA314149	CA314149 PR21E03.X	C 199	16	30.8	808	28	BH477022	BH477022 BQ6TW33TF
C 127	16	30.8	561	14	CD046845	CD046845 psHB026XM	C 200	16	30.8	809	29	BZ539581	BZ539581 OGCALN73TC
C 128	16	30.8	568	12	B1995775	B1995775 1031030A0	C 201	16	30.8	810	12	BG827959	BG827959 602752158
C 129	16	30.8	576	13	BQ694388	BQ694388 1000531.H	C 202	16	30.8	816	14	CB665452	CB665452 OSJNE4110
C 130	16	30.8	582	14	CD033238	CD033238 mgmC012XB	C 203	16	30.8	831	14	CA787598	CA787598 AGENCURT
C 131	16	30.8	585	12	BM866386	BM866386 mgc0005XJ	C 204	16	30.8	831	29	BZ539587	BZ539587 OGCALN73TM
C 132	16	30.8	588	12	BH226976	BH226976 LJNE532E	C 205	16	30.8	832	29	CNS04R9Y	CNS04R9Y Atxph1e8
C 133	16	30.8	589	14	CD042248	CD042248 psHB039XM	C 206	16	30.8	835	29	CNS04R9Y	CNS04R9Y Tetraodon
C 134	16	30.8	593	28	BH916222	BH916222 3526_1_51	C 207	16	30.8	839	29	CC443641	CC443641 pMH5C50TB
C 135	16	30.8	596	28	BH231769	BH231769 1006164C0	C 208	16	30.8	843	10	BE729005	BE729005 601562267
C 136	16	30.8	597	12	B1994012	B1994012 1031006E1	C 209	16	30.8	847	29	CNS042XU	CNS042XU Tetradon
C 137	16	30.8	598	10	BE257206	BE257206 601109572	C 210	16	30.8	850	29	BZ556298	BZ556298 pasc1-60_
C 138	16	30.8	602	14	CD208037	CD208037 HS1_36-CO	C 211	16	30.8	864	11	AY110603	AY110603 Zea_mays
C 139	16	30.8	602	10	BG116222	BG116222 602318411	C 212	16	30.8	866	11	BG110564	BG110564 602281003
C 140	16	30.8	607	29	DR6M20T	DR6M20T Dan1o.rer	C 213	16	30.8	877	29	CNS010XN	CNS010XN Anopheles
C 141	16	30.8	610	13	BQ295080	BQ295080 WHE2858.A	C 214	16	30.8	880	29	CNS010XN	CNS010XN Anopheles
C 142	16	30.8	610	28	BG472574	BG472574 602514330	C 215	16	30.8	882	29	CNS03KXN	CNS03KXN Tetradon
C 143	16	30.8	612	10	BG472574	BG472574 602514330	C 216	16	30.8	883	13	BE916224	BE916224 BX467356
C 144	16	30.8	613	28	BH416078	BH416078 1007046C0	C 217	16	30.8	884	10	BE916224	BE916224 601664076
C 145	16	30.8	616	14	CD042573	CD042573 psHB041Xr	C 218	16	30.8	896	9	AL933492	AL933492 AL933492
C 146	16	30.8	617	12	BM649582	BM649582 170006873	C 219	16	30.8	901	13	B0886962	B0886962 AGENCURT
C 147	16	30.8	618	14	CA731366	CA731366 w1d1c.pk0	C 220	16	30.8	920	29	AL930137	AL930137 NL30137
C 148	16	30.8	618	28	BH816939	BH816939 AM_Ba002	C 221	16	30.8	928	29	CC111597	CC111597 NTL_60M15
C 149	16	30.8	619	9	A1833675	A1833675 605053G05	C 222	16	30.8	935	10	BF218738	BF218738 601885106
C 150	16	30.8	620	13	BQ838365	BQ838365 WHE2909_F	C 223	16	30.8				

C 224	16	30.8	935	12	BM816946	HB106H03	C 297	15	28.8	327	10	AM923173	AM923173	DCI_49_A1
C 225	16	30.8	946	29	CC008109	PUGKD25TB	C 298	15	28.8	329	10	BE926505	BE926505	OV1_BT063
C 226	16	30.8	958	10	BE885719	BE885719	C 299	15	28.8	334	13	BU100012	BU100012	WHE3J14_B
C 227	16	30.8	975	29	CNS03G83	AL243336	C 300	15	28.8	337	14	T14455	T14455	SME5T0116_S
C 228	16	30.8	983	12	BM043376	BM043376	C 301	15	28.8	339	10	BE694037	BE694037	OV2_BT061
C 229	16	30.8	994	13	BU526783	BU526783	C 302	15	28.8	341	10	BF473422	BF473422	WHE0924_C
C 230	16	30.8	996	12	BM915392	BM915392	C 303	15	28.8	342	10	BE245637	BE245637	TCBAP1E32
C 231	16	30.8	1002	13	BM370171	BM370171	C 304	15	28.8	345	12	BM501675	BM501675	952034F05
C 232	16	30.8	1004	28	AO330215	AO330215	C 305	15	28.8	349	10	BF021209	BF021209	UX54B01_X
C 233	16	30.8	1005	29	CC436644	CC436644	C 306	15	28.8	353	14	CA433716	CA433716	UI-H-CO0-
C 234	16	30.8	1007	10	BE540641	BE540641	C 307	15	28.8	356	9	AA280978	AA280978	z997f03.r
C 235	16	30.8	1010	29	CNS0109F	AL167856	C 308	15	28.8	359	14	CNS010PO	CNS010PO	Anopheles
C 236	16	30.8	1013	28	CC008113	CC008113	C 309	15	28.8	366	14	RI9584	RI9584	yg28d07.r1
C 237	16	30.8	1025	29	BM651178	BM651178	C 310	15	28.8	368	14	BM036260	BM036260	BY036260
C 238	16	30.8	1029	29	CC280408	CC280408	C 311	15	28.8	373	13	BM0759134	BM0759134	EBma07_SQ
C 239	16	30.8	1043	29	CNS01P8N	AL134503	C 312	15	28.8	376	28	AO786035	AO786035	HS_3082_B
C 240	16	30.8	1067	13	BU172793	BU172793	C 313	15	28.8	380	10	BE496855	BE496855	WHE0761_D
C 241	16	30.8	1071	10	BE612516	BE612516	C 314	15	28.8	382	12	CB810577	CB810577	AMGNNUC1S
C 242	16	30.8	1073	10	BE846543	BE846543	C 315	15	28.8	386	9	AV641379	AV641379	AV641379
C 243	16	30.8	1091	12	BM552976	BM552976	C 316	15	28.8	390	12	BM163395	BM163395	HS_2252_A
C 244	16	30.8	1092	29	CNS05778	AL334269	C 317	15	28.8	393	9	AV546328	AV546328	AV546328
C 245	16	30.8	1098	12	BM920485	BM920485	C 318	15	28.8	396	28	AO129889	AO129889	HS_2252_A
C 246	16	30.8	1102	13	BM063824	BM063824	C 319	15	28.8	399	29	BE663329	BE663329	147072_MA
C 247	16	30.8	1120	29	CC306863	CC306863	C 320	15	28.8	401	14	W99563	W99563	T9ESTz04D
C 248	16	30.8	1146	10	BM259686	BM259686	C 321	15	28.8	406	12	BM499757	BM499757	952034F05
C 249	16	30.8	1185	12	BI957152	BI957152	C 322	15	28.8	408	9	BM223047	BM223047	AU223047
C 250	16	30.8	1201	9	AL563226	AL563226	C 323	15	28.8	410	13	BM066236	BM066236	AG-ND-166
C 251	16	30.8	1248	12	BM060068	BM060068	C 324	15	28.8	412	28	BM222604	BM222604	1006108C0
C 252	16	30.8	1255	29	AG083036	AG083036	C 325	15	28.8	415	9	AV656566	AV656566	AV656566
C 253	16	30.8	1640	29	AG099949	AG099949	C 326	15	28.8	419	9	BM404463	BM404463	AG-ND-166
C 254	16	30.8	1823	29	AG079572	AG079572	C 327	15	28.8	424	9	AI1454952	AI1454952	LD01002.5
C 255	16	30.8	1878	12	BM845876	BM845876	C 328	15	28.8	425	10	BB846562	BB846562	BB846562
C 256	16	30.8	3088	29	BZ548676	BZ548676	C 329	15	28.8	425	14	TO6839	TO6839	EST04728_Fe
C 257	16	30.8	3377	29	BZ576368	BZ576368	C 330	15	28.8	431	9	AL602345	AL602345	DKFZP868B
C 258	16	30.8	86	13	BQ757965	BQ757965	C 331	15	28.8	431	10	BF705057	BF705057	RH122_1_H
C 259	15	28.8	95	13	BM652348	BM652348	C 332	15	28.8	433	9	BM553418	BM553418	BM553418
C 260	15	28.8	112	10	BE355193	BE355193	C 333	15	28.8	434	28	BM638922	BM638922	1008021F0
C 261	15	28.8	136	12	BM299445	BM299445	C 334	15	28.8	435	10	BG355971	BG355971	EM1_18_HO
C 262	15	28.8	154	10	BM662167	BM662167	C 335	15	28.8	438	12	BI684708	BI684708	603307054
C 263	15	28.8	156	12	BM888738	BM888738	C 336	15	28.8	441	9	AV600393	AV600393	AMGNNUC1S
C 264	15	28.8	158	9	AL822647	AL822647	C 337	15	28.8	444	12	BM751097	BM751097	AV600393
C 265	15	28.8	184	12	BI017817	BI017817	C 338	15	28.8	445	9	AV952777	AV952777	AV952777
C 266	15	28.8	188	12	BM550427	BM550427	C 339	15	28.8	449	12	BM499758	BM499758	AU036472
C 267	15	28.8	197	11	CNS09AEL	BM049561	C 340	15	28.8	449	13	BU034908	BU034908	603608555
C 268	15	28.8	209	10	BF778825	BF778825	C 341	15	28.8	451	13	BU050329	BU050329	111027A0
C 269	15	28.8	217	9	AV151742	AV151742	C 342	15	28.8	452	14	CD561318	CD561318	CD561318
C 270	15	28.8	220	9	AA566913	AA566913	C 343	15	28.8	453	9	AU223112	AU223112	AU223112
C 271	15	28.8	243	12	BI056483	BI056483	C 344	15	28.8	455	13	BM074484	BM074484	946112A06
C 272	15	28.8	249	10	BE475719	BE475719	C 345	15	28.8	456	9	AA690106	AA690106	vt79C06.8
C 273	15	28.8	251	10	BE453910	BE453910	C 346	15	28.8	457	10	BE248578	BE248578	NFO09H11D
C 274	15	28.8	255	14	CA753417	CA753417	C 347	15	28.8	460	9	AI906527	AI906527	952059C05
C 275	15	28.8	260	9	AA916565	AA916565	C 348	15	28.8	461	13	BM113269	BM113269	wf28B01.X
C 276	15	28.8	267	10	BE684563	BE684563	C 349	15	28.8	461	28	AZ299641	AZ299641	RPCT_23-4
C 277	15	28.8	267	10	BE684564	BE684564	C 350	15	28.8	466	10	BE245021	BE245021	TCBAP1E27
C 278	15	28.8	267	10	BE684565	BE684565	C 351	15	28.8	466	10	BE609364	BE609364	BM195466
C 279	15	28.8	267	10	BE684570	BE684570	C 352	15	28.8	473	13	BM195466	BM195466	BM195466
C 280	15	28.8	267	10	BE721591	BE721591	C 353	15	28.8	476	9	AA227360	AA227360	zt17B04.r
C 281	15	28.8	277	10	BM521205	BM521205	C 354	15	28.8	478	9	AU995358	AU995358	AU995358
C 282	15	28.8	282	9	AM521615	AM521615	C 355	15	28.8	478	12	BI227283	BI227283	602948690
C 283	15	28.8	282	11	AK013572	AK013572	C 356	15	28.8	481	28	AO846572	AO846572	LMJAFV1_1
C 284	15	28.8	290	10	BE830574	BE830574	C 357	15	28.8	481	28	AO849021	AO849021	LMJAFV1_1
C 285	15	28.8	291	14	BY125337	BY125337	C 358	15	28.8	481	10	BE853996	BE853996	ux25405.Y
C 286	15	28.8	293	9	AI576874	AI576874	C 359	15	28.8	481	12	BI916505	BI916505	603178401
C 287	15	28.8	293	10	BM515514	BM515514	C 360	15	28.8	481	12	BI916505	BI916505	603178401
C 288	15	28.8	297	9	AV962501	AV962501	C 361	15	28.8	481	12	BI916505	BI916505	603178401
C 289	15	28.8	300	9	AL841736	AL841736	C 362	15	28.8	481	12	BI916505	BI916505	603178401
C 290	15	28.8	317	14	CD425658	CD425658	C 363	15	28.8	481	12	BI916505	BI916505	603178401
C 291	15	28.8	318	13	AL927911	AL927911	C 364	15	28.8	481	12	BI916505	BI916505	603178401
C 292	15	28.8	318	13	BU060669	BU060669	C 365	15	28.8	481	12	BI916505	BI916505	603178401
C 293	15	28.8	319	9	AA566825	AA566825	C 366	15	28.8	481	12	BI916505	BI916505	603178401
C 294	15	28.8	323	28	AZ612592	AZ612592	C 367	15	28.8	481	12	BI916505	BI916505	603178401
C 295	15	28.8	326	10	BG61863	BG61863	C 368	15	28.8	481	12	BI916505	BI916505	603178401
C 296	15	28.8	326	10	AM947629	AM947629	C 369	15	28.8	481	12	BI916505	BI916505	603178401

C 370	15	28.8	482	28	BH394006	BH394006 AC-ND-166	C 443	15	28.8	551	9	A1534769	A1534769 SD07777.5
C 371	15	28.8	483	9	A1737204	A1737204 606037D03	C 444	15	28.8	551	29	AG218247	AG218247 D960010
C 372	15	28.8	483	9	AM410695	AM410695 f607h01.x	C 445	15	28.8	551	29	AG218247	AG218247 D960010
C 373	15	28.8	483	9	AM446689	AM446689 86725 MAR	C 446	15	28.8	552	14	CA704626	CA704626 wdk1c.pk0
C 374	15	28.8	483	12	BU162647	BU162647 BU162647	C 447	15	28.8	552	14	CA704626	CA704626 wdk1c.pk0
C 375	15	28.8	489	12	BU159455	BU159455 BU159455	C 448	15	28.8	555	10	BE304855	BE304855 w1k8.pk00
C 376	15	28.8	492	28	BH638580	BH638580 1008023D0	C 449	15	28.8	557	9	A1665925	A1665925 601143730
C 377	15	28.8	493	28	AZ896229	AZ896229 RPCT-24-1	C 450	15	28.8	557	9	AM352609	AM352609 606003C09
C 378	15	28.8	494	10	BG750010	BG750010 602708766	C 451	15	28.8	557	10	BE598423	BE598423 P11 82.G0
C 379	15	28.8	494	14	CB375410	CB375410 rxd6e06.y	C 452	15	28.8	557	10	BE598423	BE598423 P11 82.G0
C 380	15	28.8	494	28	AQ222566	AQ222566 HS_3252.B	C 453	15	28.8	558	9	A1692103	A1692103 606012D05
C 381	15	28.8	496	10	BG265797	BG265797 1000081E0	C 454	15	28.8	558	10	BE397043	BE397043 601289717
C 382	15	28.8	498	9	AM249049	AM249049 2820994.5	C 455	15	28.8	558	13	BU062006	BU062006 Fgr_1.D14
C 383	15	28.8	499	12	BX253701	BX253701 BX253701	C 456	15	28.8	561	28	BH775964	BH775964 f2mb011fo
C 384	15	28.8	499	13	BM661324	BM661324 952046G06	C 457	15	28.8	564	14	CD363377	CD363377 UI-M-GLO-
C 385	15	28.8	502	9	AL036071	AL036071 DKF2P564F	C 458	15	28.8	565	12	BI420040	BI420040 LjNE5T52b
C 386	15	28.8	502	14	CD234788	CD234788 S51_18.B0	C 459	15	28.8	565	12	BI420040	BI420040 LjNE5T52b
C 387	15	28.8	502	28	BH394468	BH394468 AG-ND-167	C 460	15	28.8	566	12	BM175399	BM175399 TGESt2yb2
C 388	15	28.8	502	29	CC162677	CC162677 1183903.b	C 461	15	28.8	566	14	R90100	R90100 16455.Lambd
C 389	15	28.8	503	10	BE802749	BE802749 s7c57b01.y	C 462	15	28.8	567	13	BU788556	BU788556 1184b06.y
C 390	15	28.8	505	10	AZ142109	AZ142109 SP_0019.B	C 463	15	28.8	570	12	BM344325	BM344325 r149e04.y
C 391	15	28.8	505	14	CB853665	CB853665 UI-CF-FNO	C 464	15	28.8	571	10	BE872294	BE872294 601446314
C 392	15	28.8	507	12	BI560563	BI560563 BI560563	C 465	15	28.8	573	9	AM247583	AM247583 2820132.5
C 393	15	28.8	507	13	BO303823	BO303823 RCT-B7025	C 466	15	28.8	573	10	BE583249	BE583249 8-4E-MY.P
C 394	15	28.8	508	28	AQ815455	AQ815455 HS_5291.B	C 467	15	28.8	573	12	BI565264	BI565264 RH62992.5
C 395	15	28.8	509	9	AM759943	AM759943 s156a09.y	C 468	15	28.8	573	12	BI616411	BI616411 RH46351.5
C 396	15	28.8	510	14	CA702122	CA702122 wdk1c.pk0	C 469	15	28.8	573	12	BM704551	BM704551 UI-E-CO1-
C 397	15	28.8	511	12	BI811275	BI811275 M002P12.O	C 470	15	28.8	573	13	BO908229	BO908229 T004G08.O
C 398	15	28.8	514	9	AM705405	AM705405 s1k4e09.y	C 471	15	28.8	574	10	BE396864	BE396864 601289568
C 399	15	28.8	514	10	BF551808	BF551808 UI-R-C1-K	C 472	15	28.8	574	12	BI626692	BI626692 RH67705.5
C 400	15	28.8	515	12	BI995829	BI995829 1031030F0	C 473	15	28.8	574	12	BM791120	BM791120 K-EST0071
C 401	15	28.8	515	14	CD425973	CD425973 S41_15.A0	C 474	15	28.8	575	12	BM654181	BM654181 170068793
C 402	15	28.8	516	12	BU560953	BU560953 BU560953	C 475	15	28.8	579	10	BG524336	BG524336 41.50.Ste
C 403	15	28.8	516	12	BP124297	BP124297 BP124297	C 476	15	28.8	580	9	A1665997	A1665997 606001H06
C 404	15	28.8	516	28	BQ048590	BQ048590 952023D12	C 477	15	28.8	582	12	BI667915	BI667915 603292813
C 405	15	28.8	516	28	AZ879610	AZ879610 RPCT-23-1	C 478	15	28.8	582	28	AQ521085	AQ521085 HS_5212.B
C 406	15	28.8	518	10	BG448022	BG448022 NF105A02E	C 479	15	28.8	585	10	BE297319	BE297319 601177622
C 407	15	28.8	522	9	AM306828	AM306828 s149b06.y	C 480	15	28.8	586	12	BI571111	BI571111 BU157111
C 408	15	28.8	522	12	BU293126	BU293126 BU293126	C 481	15	28.8	586	12	BI571111	BI571111 BU157111
C 409	15	28.8	524	28	AZ612852	AZ612852 1M0439M23	C 482	15	28.8	587	9	A1534884	A1534884 SD01150.5
C 410	15	28.8	524	28	BE666673	BE666673 uW03111.x	C 483	15	28.8	587	10	BE998314	BE998314 EST430037
C 411	15	28.8	525	12	BI556115	BI556115 BI556115	C 484	15	28.8	587	11	AI106359	AI106359 Zee mayB
C 412	15	28.8	526	10	BE666673	BE666673 uW03111.x	C 485	15	28.8	588	12	BI316065	BI316065 BU316065
C 413	15	28.8	527	28	AQ619403	AQ619403 HS_5177.B	C 486	15	28.8	590	9	AM453151	AM453151 660031H05
C 414	15	28.8	528	9	A1770925	A1770925 606063C12	C 487	15	28.8	592	10	BE513360	BE513360 601318568
C 415	15	28.8	528	12	BM259530	BM259530 952010G02	C 488	15	28.8	592	12	BU293896	BU293896 BI293896
C 416	15	28.8	528	14	BU049431	BU049431 1111008A0	C 489	15	28.8	594	9	AV983043	AV983043 AV983043
C 417	15	28.8	532	12	BI560252	BI560252 BI560252	C 490	15	28.8	595	10	BG525187	BG525187 45--44.Ste
C 418	15	28.8	533	13	BX273896	BX273896 BX273896	C 491	15	28.8	596	28	AQ333696	AQ333696 HS_5013.A
C 419	15	28.8	533	14	CA443265	CA443265 UI-H-DH1-	C 492	15	28.8	597	28	AZ432687	AZ432687 1M0218O12
C 420	15	28.8	535	10	BG051713	BG051713 FM1_60.DO	C 493	15	28.8	597	28	AZ432687	AZ432687 1M0218O12
C 421	15	28.8	536	9	AM528987	AM528987 AVS28987	C 494	15	28.8	600	10	AM963089	AM963089 EST375162
C 422	15	28.8	536	9	AM107783	AM107783 u193G06.y	C 495	15	28.8	600	12	BG803759	BG803759 0241-49.M
C 423	15	28.8	538	13	BX273897	BX273897 BX273897	C 496	15	28.8	600	12	BI987526	BI987526 3202-64.M
C 424	15	28.8	538	14	CD032498	CD032498 mgmct009XU	C 497	15	28.8	600	12	BI987526	BI987526 3202-64.M
C 425	15	28.8	538	29	TA47H09P	TA47H09P TA47H09P	C 498	15	28.8	600	12	BM487480	BM487480 pmm2n.pk0
C 426	15	28.8	539	28	AZ158946	AZ158946 SP_0061.B	C 499	15	28.8	601	14	CA527134	CA527134 8032-37.M
C 427	15	28.8	540	12	BE269499	BE269499 601184702	C 500	15	28.8	601	10	BE295991	BE295991 601174290
C 428	15	28.8	540	12	BM420007	BM420007 R022B02.O	C 501	15	28.8	601	12	AQ329955	AQ329955 BI299955
C 429	15	28.8	540	28	AQ510277	AQ510277 hxbxb00951	C 502	15	28.8	601	12	AQ329955	AQ329955 BI299955
C 430	15	28.8	541	10	BG706613	BG706613 602674103	C 503	15	28.8	602	13	BU050747	BU050747 1111053A0
C 431	15	28.8	542	10	BF565902	BF565902 602185849	C 504	15	28.8	603	9	AJ273037	AJ273037 AJ273037
C 432	15	28.8	542	12	BM420939	BM420939 U021E10.O	C 505	15	28.8	603	10	BF656768	BF656768 OV2_21.F0
C 433	15	28.8	543	12	BU123796	BU123796 603148780	C 506	15	28.8	604	14	CB825074	CB825074 r3J1f0J.y
C 434	15	28.8	543	12	BU196790	BU196790 BU196790	C 507	15	28.8	604	12	BM603493	BM603493 170006870
C 435	15	28.8	544	9	AU023795	AU023795 AU023795	C 508	15	28.8	604	13	BX275967	BX275967 BX275967
C 436	15	28.8	545	10	BF777330	BF777330 NXSI_067	C 509	15	28.8	604	28	AQ848444	AQ848444 LMAJFV1_1
C 437	15	28.8	547	12	BM379527	BM379527 M8ST506.G	C 510	15	28.8	605	14	CB279297	CB279297 r34J2E12.y
C 438	15	28.8	547	13	BU476697	BU476697 603844546	C 511	15	28.8	605	14	CB682706	CB682706 OSJNE10F
C 439	15	28.8	547	13	BU978892	BU978892 HA14J03x	C 512	15	28.8	605	29	B2363275	B2363275 1d83d03.G
C 440	15	28.8	549	12	BU554357	BU554357 BU554357	C 513	15	28.8	606	9	AV825339	AV825339 AV825339
C 441	15	28.8	549	29	CC043182	CC043182 3591.1.15	C 514	15	28.8	607	14	CB452514	CB452514 707383.MA
C 442	15	28.8	550	12	BM276772	BM276772 952010G02	C 515	15	28.8	608	13	BU655505	BU655505 1112120E0

C 516	15	28.8	608 14	CB458133	716270 MA	589	15	28.8	672 12	BM596911	BM596911 170006689
C 517	15	28.8	609 9	AV974119	AV974119	590	15	28.8	673 12	HSM070316	BX480309 Home sapi
C 518	15	28.8	610 29	BZ404096	BZ404096	591	15	28.8	673 13	BM046590	BM046590 BM046590
C 519	15	28.8	615 9	AU082980	AU082980	592	15	28.8	673 29	BX208155	BX208155 1d36d06.b
C 520	15	28.8	616 9	AL819053	AL819053	593	15	28.8	674 29	BX132713	BX132713 Danilo rer
C 521	15	28.8	618 10	BE270496	BE270496	594	15	28.8	675 10	CG705103	CG705103 602688055
C 522	15	28.8	619 13	BU632191	BU318694 603515403	595	15	28.8	676 9	A1294378	A1294378 LP07737.5
C 523	15	28.8	621 10	CA619058	CA619058	596	15	28.8	677 10	BE386173	BE386173 601274421
C 524	15	28.8	620 10	CG178485	CG178485	597	15	28.8	678 9	AW584626	AW584626 N2107116
C 525	15	28.8	620 12	BU235151	BU235151	598	15	28.8	678 12	B1733943	B1733943 603355663
C 526	15	28.8	622 12	BU280185	BU280185	599	15	28.8	680 12	B1954394	B1954394 HVSEMem01
C 527	15	28.8	622 13	BU883131	BU883131	600	15	28.8	681 14	BY179362	BY179362 BY179362
C 528	15	28.8	623 10	BF482456	BF482456	601	15	28.8	682 13	BE123985	BE123985 Fgr. H H23
C 529	15	28.8	623 10	BF483216	BF483216	602	15	28.8	682 10	BE123985	BE123985 Fgr. H H23
C 530	15	28.8	624 13	BQ295733	BQ295733	603	15	28.8	688 29	BZ328233	BZ328233 1d36d06.g
C 531	15	28.8	625 12	B1391033	B1391033	604	15	28.8	689 12	B1754206	B1754206 603025501
C 533	15	28.8	625 13	BQ169269	BQ169269	605	15	28.8	690 10	BG522673	BG522673 24-15 Ste
C 534	15	28.8	628 14	CA755709	CA755709	606	15	28.8	692 10	BG086072	BG086072 WHE3574 E
C 535	15	28.8	629 12	BU203491	BU203491	607	15	28.8	692 10	BG120794	BG120794 602348557
C 536	15	28.8	630 13	BU397278	BU397278	608	15	28.8	693 10	BG068783	BG068783 H3069C06-
C 537	15	28.8	632 10	BF587572	BF587572	609	15	28.8	694 13	BU371405	BU371405 603595679
C 538	15	28.8	632 10	BF587572	BF587572	610	15	28.8	694 28	BH656620	BH656620 BOWM149TF
C 539	15	28.8	633 9	A1677240	A1677240	611	15	28.8	695 10	BE388641	BE388641 601283478
C 540	15	28.8	633 9	BZ544058	BZ544058	612	15	28.8	695 29	CNS06DM6	AL393988 T3 end of
C 541	15	28.8	635 29	BZ544058	BZ544058	613	15	28.8	696 10	CG081797	CG081797 H3069C06-
C 542	15	28.8	639 12	BH601368	BH601368	614	15	28.8	696 10	BG686954	BG686954 602650948
C 543	15	28.8	639 12	BM301007	BM301007	615	15	28.8	698 10	BF683235	BF683235 602139253
C 544	15	28.8	641 9	AM584471	AM584471	616	15	28.8	698 12	BG914072	BG914072 602805586
C 545	15	28.8	641 10	BF686563	BF686563	617	15	28.8	699 13	BG653832	BG653832 1112109C1
C 546	15	28.8	644 10	BE296757	BE296757	618	15	28.8	700 10	BG463806	BG463806 EM1 51 BO
C 547	15	28.8	644 10	BE296757	BE296757	619	15	28.8	701 10	BCG701097	BCG701097 602681529
C 548	15	28.8	644 10	BE296757	BE296757	620	15	28.8	702 13	BX254177	BX254177 BX254177
C 549	15	28.8	644 10	BE296757	BE296757	621	15	28.8	702 13	BX254177	BX254177 BX254177
C 550	15	28.8	644 10	BE296757	BE296757	622	15	28.8	704 12	B1483409	B1483409 RE66306.5
C 551	15	28.8	645 10	BF685440	BF685440	623	15	28.8	710 13	BU273877	BU273877 603375275
C 552	15	28.8	646 10	BE867889	BE867889	624	15	28.8	711 29	BZ639904	BZ639904 OGA0254TC
C 553	15	28.8	646 10	BE867889	BE867889	625	15	28.8	712 10	BE965750	BE965750 601659514
C 554	15	28.8	649 12	BI529212	BI529212	626	15	28.8	712 13	BQ999753	BQ999753 Fgr. 1 D14
C 555	15	28.8	649 12	BI529212	BI529212	627	15	28.8	713 13	BQ999753	BQ999753 OGA0254TC
C 556	15	28.8	650 10	BG521416	BG521416	628	15	28.8	714 10	BE270110	BE270110 601185913
C 557	15	28.8	651 13	BU380081	BU380081	629	15	28.8	715 10	BG523071	BG523071 28-57 Ste
C 558	15	28.8	654 12	BP004222	BP004222	630	15	28.8	716 13	BU361615	BU361615 603790064
C 559	15	28.8	655 12	BP004222	BP004222	631	15	28.8	720 10	BG476490	BG476490 602523585
C 560	15	28.8	655 14	CB378796	CB378796	632	15	28.8	720 29	CNS01EPM	AL10982 ANOPH168
C 561	15	28.8	656 10	BF648356	BF648356	633	15	28.8	722 13	BM240142	BM240142 BM240142
C 562	15	28.8	657 10	BF648356	BF648356	634	15	28.8	724 13	BU389597	BU389597 603807307
C 563	15	28.8	657 29	BZ566247	BZ566247	635	15	28.8	726 10	CC338221	CC338221 OGB1878TV
C 564	15	28.8	658 10	BG719975	BG719975	636	15	28.8	727 13	BU112900	BU112900 603130144
C 565	15	28.8	658 13	BU059936	BU059936	637	15	28.8	727 29	BZ827095	BZ827095 602522025
C 566	15	28.8	658 13	BU059936	BU059936	638	15	28.8	728 10	BE294708	BE294708 601179304
C 567	15	28.8	658 13	BU059936	BU059936	639	15	28.8	728 10	BF304732	BF304732 60188304
C 568	15	28.8	658 28	BH816398	BH816398	640	15	28.8	728 29	BZ457717	BZ457717 BONR281TR
C 569	15	28.8	658 29	CC107164	CC107164	641	15	28.8	728 29	BZ681486	BZ681486 PUBA168TD
C 570	15	28.8	660 13	BQ407422	BQ407422	642	15	28.8	730 29	BZ643745	BZ643745 OGBA75TM
C 571	15	28.8	660 13	BQ407422	BQ407422	643	15	28.8	730 29	CNS045XM	AL305579 Tetracodon
C 572	15	28.8	661 13	BQ8339036	BQ8339036	644	15	28.8	732 29	BZ739448	BZ739448 OGBA45TM
C 573	15	28.8	661 29	BZ351810	BZ351810	645	15	28.8	733 12	B1116859	B1116859 602868773
C 574	15	28.8	662 13	BU340377	BU340377	646	15	28.8	734 10	BG758517	BG758517 602712729
C 575	15	28.8	663 14	CA680547	CA680547	647	15	28.8	736 9	AJ558966	AJ558966 AJ558966
C 576	15	28.8	663 14	CA680547	CA680547	648	15	28.8	736 9	AJ558966	AJ558966 AJ558966
C 577	15	28.8	664 13	BO407491	BO407491	649	15	28.8	737 13	BX084601	BX084601 BX084601
C 578	15	28.8	664 13	BO407491	BO407491	650	15	28.8	738 29	BZ73567	BZ73567 OGFCK74TM
C 579	15	28.8	666 14	CD212192	CD212192	651	15	28.8	739 14	CB989245	CB989245 ACENECOURT
C 580	15	28.8	666 14	CD212192	CD212192	652	15	28.8	740 12	BM011746	BM011746 603636115
C 581	15	28.8	667 9	AU128057	AU128057	653	15	28.8	744 13	BU468227	BU468227 603371182
C 582	15	28.8	667 12	BU575521	BU575521	654	15	28.8	745 12	B1559838	B1559838 603252753
C 583	15	28.8	667 12	BU575521	BU575521	655	15	28.8	747 12	BM051889	BM051889 603628765
C 584	15	28.8	668 12	BM425910	BM425910	656	15	28.8	748 13	BX375170	BX375170 BX375170
C 585	15	28.8	670 13	BU061912	BU061912	657	15	28.8	750 9	A1534941	A1534941 SD01232.5
C 586	15	28.8	670 13	BM275510	BM275510	658	15	28.8	751 29	BZ735658	BZ735658 OGFCK74TC
C 587	15	28.8	672 12	BG779584	BG779584	659	15	28.8	753 9	AU142422	AU142422 AU142422
C 588	15	28.8	672 12	BG779584	BG779584	661	15	28.8	754 10	BM438057	BM438057 602490319

662	15	28.8	758	14	CD388217	AGENCOURT	735	15	28.8	865	29	B2404884	B2404884	OGACAK1TC
C 663	15	28.8	760	12	B1559353	AGENCOURT	C 736	15	28.8	867	14	CB988849	CB988849	AGENCOURT
C 664	15	28.8	761	9	AV384007	AV384007	C 737	15	28.8	871	13	BU462228	BU462228	AGENCOURT
C 665	15	28.8	761	10	BF983495	602307149	C 738	15	28.8	875	13	BU542338	BU542338	AGENCOURT
C 666	15	28.8	765	12	BC760180	602307149	C 739	15	28.8	877	10	BE868031	BE868031	AGENCOURT
C 667	15	28.8	765	29	CNS0405X	602307149	C 740	15	28.8	879	10	BU166721	BU166721	AGENCOURT
C 668	15	28.8	766	13	BQ788995	WME4156_C	C 741	15	28.8	880	12	B1915536	B1915536	AGENCOURT
C 669	15	28.8	767	12	BQ772615	602720724	C 742	15	28.8	880	13	BQ212514	BQ212514	AGENCOURT
C 670	15	28.8	767	12	B1763787	602720724	C 743	15	28.8	881	10	BG446687	BG446687	GA_Fa002
C 671	15	28.8	767	28	AZ212615	602720724	C 744	15	28.8	885	10	BC326741	BC326741	602720724
C 672	15	28.8	768	12	BM077278	602720724	C 745	15	28.8	889	13	BU541159	BU541159	AGENCOURT
C 673	15	28.8	769	29	CC340130	602720724	C 746	15	28.8	889	29	B2571332	B2571332	AGENCOURT
C 674	15	28.8	770	12	BI227337	602948960	C 747	15	28.8	890	10	BU359828	BU359828	AGENCOURT
C 675	15	28.8	771	11	CNS08J43	602948960	C 748	15	28.8	892	12	BI603374	BI603374	602948960
C 676	15	28.8	771	14	CB655905	OSJNEC09L	C 749	15	28.8	893	29	B2638590	B2638590	OGCBV82TC
C 677	15	28.8	772	29	B2604461	602948960	C 750	15	28.8	895	10	BC701833	BC701833	602948960
C 678	15	28.8	773	29	B2739438	602948960	C 751	15	28.8	896	29	CC091647	CC091647	CSU-K33r
C 679	15	28.8	776	29	CNS04A02P	602948960	C 752	15	28.8	896	29	CNS04222	CNS04222	Tetraodon
C 680	15	28.8	777	14	CB940761	IPCGX14	C 753	15	28.8	904	10	BG438198	BG438198	602948960
C 681	15	28.8	777	29	B2657183	OGCBX12TC	C 754	15	28.8	905	13	BU162710	BU162710	AGENCOURT
C 682	15	28.8	778	10	BG524305	41-21_Ste	C 755	15	28.8	906	10	BF783988	BF783988	602107831
C 683	15	28.8	778	13	BU281449	603577293	C 756	15	28.8	906	10	BE893084	BE893084	601436573
C 684	15	28.8	779	13	BU284773	603602362	C 757	15	28.8	908	29	CNS01KX8	AL146693	Anophe1eg
C 685	15	28.8	780	14	CD109543	AGENCOURT	C 758	15	28.8	909	12	BG826537	BG826537	602750221
C 686	15	28.8	781	12	BM012936	603642904	C 759	15	28.8	910	10	BG340603	BG340603	602462174
C 687	15	28.8	782	12	BM016575	603642904	C 760	15	28.8	910	13	BQ949736	BQ949736	AGENCOURT
C 688	15	28.8	784	28	BH542166	BOHRM18TR	C 761	15	28.8	910	13	BU169324	BU169324	AGENCOURT
C 689	15	28.8	786	12	B1460896	603207114	C 762	15	28.8	910	13	BC335940	BC335940	60335940
C 690	15	28.8	787	29	B2639912	OGAC0254TM	C 763	15	28.8	912	13	BQ954945	BQ954945	AGENCOURT
C 691	15	28.8	790	12	BI948694	603207114	C 764	15	28.8	913	10	BF678272	BF678272	602086184
C 692	15	28.8	793	10	BE544179	601076732	C 765	15	28.8	913	12	BI913314	BI913314	603180079
C 693	15	28.8	793	29	B2570319	603055257	C 766	15	28.8	914	13	BO431719	BO431719	AGENCOURT
C 694	15	28.8	794	12	BI771858	603055257	C 767	15	28.8	914	13	BO883362	BO883362	AGENCOURT
C 695	15	28.8	797	14	CAS88543	hab65b03	C 768	15	28.8	914	13	BC369513	BC369513	60369513
C 696	15	28.8	798	10	BG170643	602323508	C 769	15	28.8	915	28	AZ136822	SP_0173_A	
C 697	15	28.8	799	10	BG432825	602449390	C 770	15	28.8	915	28	BU146443	BU146443	AGENCOURT
C 698	15	28.8	802	10	BE413693	602449390	C 771	15	28.8	916	13	BO883186	BO883186	AGENCOURT
C 699	15	28.8	806	13	BU477167	603471324	C 772	15	28.8	917	13	BU468770	BU468770	603370908
C 700	15	28.8	807	29	CNS01MSX	602449390	C 773	15	28.8	918	13	CA454895	CA454895	AGENCOURT
C 701	15	28.8	808	12	BG828556	602449390	C 774	15	28.8	918	14	BF784287	BF784287	602108157
C 702	15	28.8	810	10	BG332393	602449390	C 775	15	28.8	919	10	BI912680	BI912680	603176328
C 703	15	28.8	811	29	B2643738	602449390	C 776	15	28.8	920	12	BI156759	BI156759	603254110
C 704	15	28.8	811	29	CC407611	602449390	C 777	15	28.8	920	12	BI115784	BI115784	602866238
C 705	15	28.8	816	13	BU284285	603864845	C 778	15	28.8	921	13	BO886477	BO886477	AGENCOURT
C 706	15	28.8	816	29	B2451118	BOAN24TF	C 779	15	28.8	921	13	BU302820	BU302820	603737821
C 707	15	28.8	818	10	BG701407	602682616	C 780	15	28.8	921	13	BU135439	BU135439	603122502
C 708	15	28.8	818	13	BX084602	602682616	C 781	15	28.8	922	13	BQ641379	BQ641379	AGENCOURT
C 709	15	28.8	821	10	BG387421	602456187	C 782	15	28.8	926	13	BG186178	BG186178	RST5026_A
C 710	15	28.8	821	10	BC754095	602726186	C 783	15	28.8	927	28	AQ330312	AQ330312	nxh0046N
C 711	15	28.8	822	12	BG822070	602726186	C 784	15	28.8	927	28	CNS032K0	AL224937	Tetraodon
C 712	15	28.8	823	13	BU167679	AGENCOURT	C 785	15	28.8	929	29	BU151326	BU151326	AGENCOURT
C 713	15	28.8	823	28	BH701245	BOMN169TR	C 786	15	28.8	932	13	BU387083	BU387083	60387083
C 714	15	28.8	824	28	BI914256	603180643	C 787	15	28.8	935	10	BC758303	BC758303	602712487
C 715	15	28.8	828	10	BG436766	602489189	C 788	15	28.8	936	10	BG748019	BG748019	602705509
C 716	15	28.8	829	29	BZ981395	602489189	C 789	15	28.8	936	12	BI100131	BI100131	602884821
C 717	15	28.8	830	12	BI114661	PUC1006TD	C 790	15	28.8	940	10	BE743497	BE743497	601573382
C 718	15	28.8	833	12	BZ570407	602861867	C 791	15	28.8	940	10	BU905050	BU905050	AGENCOURT
C 719	15	28.8	834	12	BI462493	603203988	C 792	15	28.8	945	13	BU899981	BU899981	AGENCOURT
C 720	15	28.8	838	10	BF036203	601458275	C 793	15	28.8	945	13	BU899981	BU899981	AGENCOURT
C 721	15	28.8	840	29	BZ981393	PUC1006TD	C 794	15	28.8	947	13	BU899981	BU899981	AGENCOURT
C 722	15	28.8	841	14	BI753010	603025791	C 795	15	28.8	947	13	BU899981	BU899981	AGENCOURT
C 723	15	28.8	841	14	CB939746	603025791	C 796	15	28.8	947	13	BU899981	BU899981	AGENCOURT
C 724	15	28.8	845	10	BG188515	603025791	C 797	15	28.8	949	10	BF785750	BF785750	602112650
C 725	15	28.8	846	14	CB656014	OSJNEC09O	C 798	15	28.8	953	13	BO943589	BO943589	AGENCOURT
C 726	15	28.8	848	13	BU384305	603861191	C 799	15	28.8	962	29	CNS04SGO	BE791755	601581578
C 727	15	28.8	851	13	BU265224	603861191	C 800	15	28.8	966	10	BI083176	BI083176	602875267
C 728	15	28.8	851	29	AG111435	603861191	C 801	15	28.8	966	12	W11634	W11634	ma90h04_r1
C 729	15	28.8	853	10	BC339313	602436957	C 802	15	28.8	972	29	CNS0253K	AL151577	Tetraodon
C 730	15	28.8	854	13	BU599673	602436957	C 803	15	28.8	973	13	BU505413	BU505413	AGENCOURT
C 731	15	28.8	858	12	BI913775	603180148	C 804	15	28.8	975	29	CNS01UGS	AL167797	Tetraodon
C 732	15	28.8	858	14	CB988154	AGENCOURT	C 805	15	28.8	977	10	BE887958	BE887958	601511453
C 733	15	28.8	860	10	BG183956	RST2870_A	C 806	15	28.8	985	10	BE562273	BE562273	601344749
C 734	15	28.8	861	29	CNS07603	601344749	C 807	15	28.8	985	10	BE562273	BE562273	601344749

C 808	15	28.8	985	29	CNS06RMU	AL412156 T7 end of	C 881	15	28.8	1615	12	B1519263	B1519253 603062017
C 809	15	28.8	986	29	CNS03Y4A	AL265883 Tetraodon	C 882	15	28.8	1680	10	BF308965	BF308965 601889636
C 810	15	28.8	989	10	BE731209	BE731209 601567292	C 883	15	28.8	1761	10	BF035285	BF035285 601457171
C 811	15	28.8	993	13	BU304812	BU304812 603758440	C 884	15	28.8	1786	12	BC843868	BC843868 102400400
C 812	15	28.8	996	29	AC082404	AC082404 Pan trog1	C 885	15	28.8	1816	29	AC070404	AC070404 Pan trog1
C 813	15	28.8	997	9	AL514952	AL514952 AL514952	C 886	15	28.8	1839	12	BF527004	BF527004 602070285
C 814	15	28.8	997	29	BZ578215	BZ578215 msh2 5751	C 887	15	28.8	1854	10	BF527004	BF527004 602070285
C 815	15	28.8	997	29	CNS06QFE	AL410592 T7 end of	C 888	15	28.8	1894	29	AC070136	AC070136 Pan trog1
C 816	15	28.8	999	13	BO068056	BO068056 AGENCOURT	C 889	15	28.8	2078	12	BC843946	BC843946 102400400
C 817	15	28.8	1008	10	BE779062	BE779062 601464830	C 890	15	28.8	2229	11	A1105265	A1105265 Zea mays
C 818	15	28.8	1009	13	BA404436	BA404436 BX404436	C 891	15	28.8	2423	11	BC033173	BC033173 Homo sapi
C 819	15	28.8	1010	10	BG297512	BG297512 602395675	C 892	15	28.8	2445	11	AK036959	AK036959 Mus muscu
C 820	15	28.8	1022	13	BQ677452	BQ677452 AGENCOURT	C 893	15	28.8	2936	11	BC025272	BC025272 Homo sapi
C 821	15	28.8	1022	13	BX458376	BX458376 BX458376	C 894	15	28.8	3968	11	AK032050	AK032050 Mus muscu
C 822	15	28.8	1027	29	CNS02R02	AL209996 Tetraodon	C 895	15	28.8	61	9	AA665790	AA665790 ag70110.s
C 823	15	28.8	1031	29	AC084516	AC084516 Pan trog1	C 896	15	28.8	61	29	AG219119	AG219119 Drosophi1
C 824	15	28.8	1033	29	BZ561493	BZ561493 pac62-164	C 897	15	28.8	62	29	AG218163	AG218163 Drosophi1
C 825	15	28.8	1033	29	BQ276417	BQ276417 AGENCOURT	C 898	15	28.8	93	12	BM442529	BM442529 EBAN01 SO
C 826	15	28.8	1044	13	AG132263	AG132263 Pan trog1	C 899	15	28.8	102	9	AA171162	AA171162 m643d04.r
C 827	15	28.8	1047	29	CC286362	CC286362 CH261-29N	C 900	15	28.8	104	9	AA274455	AA274455 x662907.x
C 828	15	28.8	1051	29	BO680037	BO680037 AGENCOURT	C 901	15	28.8	108	28	BH687891	BH687891 BCGXB02TF
C 829	15	28.8	1059	13	BM680037	BM680037 AGENCOURT	C 902	15	28.8	116	12	BM851790	BM851790 K-EST0132
C 830	15	28.8	1061	13	BM359981	BM359981 BX359981	C 903	15	28.8	124	29	BZ651527	BZ651527 OCAOC17TC
C 831	15	28.8	1062	13	BU459469	BU459469 603365961	C 904	15	28.8	129	9	AM386850	AM386850 CM0-PT004
C 832	15	28.8	1065	12	BM800884	BM800884 AGENCOURT	C 905	15	28.8	135	12	AV278646	AV278646 AV278646
C 833	15	28.8	1066	12	BM800884	BM800884 AGENCOURT	C 906	15	28.8	135	12	BM846197	BM846197 K-EST0125
C 834	15	28.8	1073	10	BF344843	BF344843 602014134	C 907	15	28.8	136	29	CC154824	CC154824 CSU-K34.1
C 835	15	28.8	1073	10	BC121844	BC121844 602350885	C 908	15	28.8	137	12	BM645289	BM645289 170006873
C 836	15	28.8	1076	13	BU455441	BU455441 603214562	C 909	15	28.8	138	12	B1053840	B1053840 CM3-GN033
C 837	15	28.8	1076	29	CNS038KU	AL336620 Tetraodon	C 910	15	28.8	140	9	AA087966	AA087966 m39412.r
C 838	15	28.8	1082	9	AL555364	AL555364 AL555364	C 911	15	28.8	143	13	BO471994	BO471994 HV04B19
C 839	15	28.8	1085	29	CNS02QHE	AL196331 Tetraodon	C 912	15	28.8	148	9	A1338343	A1338343 6C43609.Y
C 840	15	28.8	1090	12	BM805398	BM805398 AGENCOURT	C 913	15	28.8	151	9	A1883410	A1883410 f63107.Y
C 841	15	28.8	1097	12	BM903173	BM903173 AGENCOURT	C 914	15	28.8	152	12	B1989827	B1989827 B198927
C 842	15	28.8	1100	10	BF205607	BF205607 601685752	C 915	15	28.8	153	10	BG157934	BG157934 EML 6 C01
C 843	15	28.8	1113	29	BZ564469	BZ564469 pac62-164	C 916	15	28.8	156	9	AA111339	AA111339 m053C06.r
C 844	15	28.8	1118	12	BZ693572	BZ693572 SP_Ba003	C 917	15	28.8	165	13	BU586854	BU586854 A1_em39F
C 845	15	28.8	1123	29	B1090789	B1090789 602855790	C 918	15	28.8	172	9	AU076697	AU076697 AU076697
C 846	15	28.8	1134	29	BM544889	BM544889 AGENCOURT	C 919	15	28.8	174	11	CNS091MG	BMX06012 Single re
C 847	15	28.8	1134	29	BZ560961	BZ560961 pac62-164	C 920	15	28.8	176	9	A1616174	A1616174 614014D10
C 848	15	28.8	1147	13	BO652436	BO652436 AGENCOURT	C 921	15	28.8	178	12	B1983444	B1983444 B198344
C 849	15	28.8	1149	11	AK020131	AK020131 Mus muscu	C 922	15	28.8	183	10	BB578719	BB578719 BB578719
C 850	15	28.8	1154	12	BC844875	BC844875 102400880	C 923	15	28.8	184	10	BF293616	BF293616 WHE2158.H
C 851	15	28.8	1160	29	AG112217	AG112217 Pan trog1	C 924	15	28.8	184	14	CD024418	CD024418 NXR.V.042
C 852	15	28.8	1164	29	CC290530	CC290530 CH261-172	C 925	15	28.8	186	10	BE181587	BE181587 CM0-HT063
C 853	15	28.8	1168	10	BF302806	BF302806 602032687	C 926	15	28.8	186	28	BH229572	BH229572 100615380
C 854	15	28.8	1178	29	AC080244	AC080244 Pan trog1	C 927	15	28.8	191	28	BH639888	BH639888 1008032E1
C 855	15	28.8	1194	29	AC079780	AC079780 Pan trog1	C 928	15	28.8	195	29	BZ351146	BZ351146 hv19B05.g
C 856	15	28.8	1197	9	AL533193	AL533193 AL533193	C 929	15	28.8	197	29	P7388	AL161407 Leishmani
C 857	15	28.8	1201	9	AL523416	AL523416 AL523416	C 930	15	28.8	198	9	AA474177	AA474177 ves2909.r
C 858	15	28.8	1201	13	AL541909	AL541909 AL541909	C 931	15	28.8	200	10	BF546768	BF546768 UI-R-Cl-1
C 859	15	28.8	1201	13	BM340487	BM340487 BX340487	C 932	15	28.8	202	12	B0047896	B0047896 952045G01
C 860	15	28.8	1201	13	BM364165	BM364165 BX364165	C 933	15	28.8	202	14	L46498	L46498 BNAF1993 Mu
C 861	15	28.8	1201	13	BM402350	BM402350 BX402350	C 934	15	28.8	204	28	AZ578071	AZ578071 18e01 Sho
C 862	15	28.8	1201	13	BM416631	BM416631 BX416631	C 935	15	28.8	209	10	BF885201	BF885201 M83-TN016
C 863	15	28.8	1201	13	BM423899	BM423899 BX423899	C 936	15	28.8	210	9	AA371353	AA371353 ESTR83
C 864	15	28.8	1219	12	BM543086	BM543086 AGENCOURT	C 937	15	28.8	210	28	BH913834	BH913834 3526.1.41
C 865	15	28.8	1259	12	B1197880	B1197880 602758007	C 938	15	28.8	211	29	CC445985	CC445985 PUMW27TB
C 866	15	28.8	1274	13	BM769552	BM769552 602744449	C 939	15	28.8	213	9	AL969834	AL969834 AL969834
C 867	15	28.8	1279	13	BM516461	BM516461 AGENCOURT	C 940	15	28.8	217	12	BM700166	BM700166 UI-E-DW1-
C 868	15	28.8	1286	29	AG119709	AG119709 Pan trog1	C 941	15	28.8	218	9	AV282988	AV282988 AV282988
C 869	15	28.8	1288	29	BM624102	BM624102 602647844	C 942	15	28.8	218	9	AA066196	AA066196 687007F08
C 870	15	28.8	1310	13	BM275143	BM275143 603532348	C 943	15	28.8	222	13	BY336907	BY336907 BY336907
C 871	15	28.8	1310	29	AG048141	AG048141 Pan trog1	C 944	15	28.8	225	28	BH648323	BH648323 BOHYD72TF
C 872	15	28.8	1330	29	AG048141	AG048141 Pan trog1	C 945	15	28.8	226	9	AA938789	AA938789 SMTBADAO0
C 873	15	28.8	1337	13	BM704559	BM704559 602688781	C 946	15	28.8	227	28	AZ577678	AZ577678 qe03g07.x
C 874	15	28.8	1401	13	BM535031	BM535031 AGENCOURT	C 947	15	28.8	227	28	U61525	U61525 HU061525 He
C 875	15	28.8	1401	13	BM535031	BM535031 AGENCOURT	C 948	15	28.8	230	9	AJ285342	AJ285342 4A3B-AA1-
C 876	15	28.8	1414	12	BZ578603	BZ578603 msh2 594.	C 949	15	28.8	232	28	BH018360	BH018360 L1458K.d
C 877	15	28.8	1431	29	AG086694	AG086694 Pan trog1	C 950	15	28.8	232	10	B1054731	B1054731 MR3-GN034
C 878	15	28.8	1432	12	BM475946	BM475946 Pan trog1	C 951	15	28.8	234	10	BB578446	BB578446 BB578446
C 879	15	28.8	1457	29	AG158868	AG158868 Pan trog1	C 952	15	28.8	236	10	BM578446	BM578446 BB578446
C 880	15	28.8	1462	12	BM922456	BM922456 AGENCOURT	C 953	15	28.8	236	28	AQ078747	AQ078747 CIT-HSP-2

```
C 954 14 26.9 237 9 A1988678
C 955 14 26.9 238 10 BE506373
C 956 14 26.9 239 29 CC001754
C 957 14 26.9 241 29 B2628644
C 958 14 26.9 243 9 A1621491
C 959 14 26.9 243 13 BQ163172
C 960 14 26.9 243 13 B170223
C 961 14 26.9 245 28 BH406412
C 962 14 26.9 247 10 BG658056
C 963 14 26.9 247 13 BU546528
C 964 14 26.9 248 9 A1284046
C 965 14 26.9 248 9 AV431283
C 966 14 26.9 248 13 BY073743
C 967 14 26.9 249 28 BH832799
C 968 14 26.9 249 28 BH833764
C 969 14 26.9 250 10 BF082559
C 970 14 26.9 251 13 BU660029
C 971 14 26.9 253 13 BU585451
C 972 14 26.9 254 14 CA890022
C 973 14 26.9 256 14 CB126527
C 974 14 26.9 259 10 BE497308
C 975 14 26.9 259 12 B1543121
C 976 14 26.9 262 28 BH617680
C 977 14 26.9 266 10 BE530068
C 978 14 26.9 267 9 AM706963
C 979 14 26.9 267 10 BG383097
C 980 14 26.9 269 12 BJ300206
C 981 14 26.9 272 9 AA597456
C 982 14 26.9 272 9 AM000380
C 983 14 26.9 276 10 BB242079
C 984 14 26.9 277 10 BB862181
C 985 14 26.9 277 13 BY680653
C 986 14 26.9 279 9 AV242658
C 987 14 26.9 280 12 BM745914
C 988 14 26.9 281 9 AA093779
C 989 14 26.9 282 13 BY191847
C 990 14 26.9 282 14 AZ351662
C 991 14 26.9 282 28 BJ184962
C 992 14 26.9 285 9 AJ285427
C 993 14 26.9 285 10 AM862388
C 994 14 26.9 285 13 BY237153
C 995 14 26.9 285 28 BH870227
C 996 14 26.9 286 14 CB346362
C 997 14 26.9 291 13 BQ345037
C 998 14 26.9 291 14 CD477528
C 999 14 26.9 292 12 BG933341
1000 14 26.9 292 12 BG933341
```

ALIGNMENTS

```
RESULT 1
LOCUS BF251279 395 bp mRNA linear EST 15-NOV-2001
DEFINITION BF251279 Coccidioides immitis spherule cDNA library Coccidioides
immitis cDNA clone CIAA11 5' sequence, mRNA sequence.
ACCESSION BF251279
VERSION BF251279.1 GI:16931422
KEYWORDS EST.
SOURCE Coccidioides immitis
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Omygenales; mitosporic Omygenales; Coccidioides.
REFERENCE 1 (bases 1 to 395)
AUTHORS Gardner M.J. and Kirkland, T.
TITLE Generation of ESTs from Coccidioides immitis spherule cDNA library
JOURNAL unpublished
COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
```

```
Fax: 301 838 0208
Email: gardner@igir.org.
Location/Qualifiers
FEATURES
Source 1..395
/organism="Coccidioides immitis"
/mol_type="mRNA"
/db_xref="taxon:5501"
/dev_stage="spherule"
/lab_host="SOLR"
/clone_11b="Coccidioides immitis spherule cDNA library"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 78 a 133 c 78 g 106 t
ORIGIN
Query Match 100.0%; Score 52; DB 10; Length 395;
Best Local Similarity 100.0%; Pred. No. 3.1e-17;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCAGTTCTCTACGCTCTCATGCTCTGTCGTCGCGGCTCGGCACTG 52
Db 142 ATGCAGTTCTCTACGCTCTCATGCTCTGTCGTCGCGGCTCGGCACTG 193
```

```
RESULT 2
LOCUS BF252856 401 bp mRNA linear EST 15-NOV-2001
DEFINITION BF252856 Coccidioides immitis spherule cDNA library Coccidioides
immitis cDNA clone CIAA870 5' sequence, mRNA sequence.
ACCESSION BF252856
VERSION BF252856.1 GI:16932999
KEYWORDS EST.
SOURCE Coccidioides immitis
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Omygenales; mitosporic Omygenales; Coccidioides.
REFERENCE 1 (bases 1 to 401)
AUTHORS Gardner M.J. and Kirkland, T.
TITLE Generation of ESTs from Coccidioides immitis spherule cDNA library
JOURNAL unpublished
COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@igir.org.
Location/Qualifiers
FEATURES
Source 1..401
/organism="Coccidioides immitis"
/mol_type="mRNA"
/db_xref="taxon:5501"
/clone="CIAA870"
/dev_stage="spherule"
/lab_host="SOLR"
/clone_11b="Coccidioides immitis spherule cDNA library"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 77 a 149 c 89 g 86 t
ORIGIN
Query Match 100.0%; Score 52; DB 10; Length 401;
Best Local Similarity 100.0%; Pred. No. 3.1e-17;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGCAGTTCTCTACGCTCTCATGCTCTGTCGTCGCGGCTCGGCACTG 52
Db 59 ATGCAGTTCTCTACGCTCTCATGCTCTGTCGTCGCGGCTCGGCACTG 110
RESULT 3
BF251334
```

LOCUS BF251334 531 bp mRNA linear EST 15-NOV-2001
DEFINITION EST148594 Coccidioides immitis spherule cDNA library Coccidioides
immitis cDNA clone CIAAF68 5' sequence, mRNA sequence.
ACCESSION BF251334
VERSION BF251334.1 GI:16931477
KEYWORDS Coccidioides immitis
SOURCE Coccidioides immitis
ORGANISM Coccidioides immitis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE 1 (bases 1 to 531)
AUTHORS Gardner, M.J. and Kirkland, T.
TITLE Generation of ESTs from Coccidioides immitis spherule cDNA library
JOURNAL Unpublished
COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org.
Location/Qualifiers
1..531
/organism="Coccidioides immitis"
/mol_type="mRNA"
/db_xref="taxon:5501"
/clone="CIAAF68"
/dev_stage="spherule"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 107 a 190 c 115 g 119 t
ORIGIN
Query Match 100.0%; Score 52; DB 10; Length 531;
Best Local Similarity 100.0%; Pred. No. 3.3e-17;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAGTTCTCTCAGCGCTCATCGCTCTGCTGCGTCCGCGCTCCGCACTG 52
Db 137 ATGCAGTTCTCTCAGCGCTCATCGCTCTGCTGCGTCCGCGCTCCGCACTG 188

RESULT 4
BF252723 557 bp mRNA linear EST 15-NOV-2001
LOCUS EST149986 Coccidioides immitis spherule cDNA library Coccidioides
immitis cDNA clone CIABA21 5' sequence, mRNA sequence.
ACCESSION BF252723
VERSION BF252723.1 GI:16932866
KEYWORDS EST.
SOURCE Coccidioides immitis
ORGANISM Coccidioides immitis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE 1 (bases 1 to 557)
AUTHORS Gardner, M.J. and Kirkland, T.
TITLE Generation of ESTs from Coccidioides immitis spherule cDNA library
JOURNAL Unpublished
COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org.
Location/Qualifiers
1..557
/organism="Coccidioides immitis"
/mol_type="mRNA"
/db_xref="taxon:5501"
/clone="CIABA21"

FEATURES
source

/dev_stage="spherule"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 110 a 197 c 117 g 133 t
ORIGIN
Query Match 100.0%; Score 52; DB 10; Length 557;
Best Local Similarity 100.0%; Pred. No. 3.3e-17;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAGTTCTCTCAGCGCTCATCGCTCTGCTGCGTCCGCGCTCCGCACTG 52
Db 161 ATGCAGTTCTCTCAGCGCTCATCGCTCTGCTGCGTCCGCGCTCCGCACTG 212

RESULT 5
BF251103 737 bp mRNA linear EST 15-NOV-2001
LOCUS EST148360 Coccidioides immitis spherule cDNA library Coccidioides
immitis cDNA clone CIAAC67 5' sequence, mRNA sequence.
ACCESSION BF251103
VERSION BF251103.1 GI:16931246
KEYWORDS EST.
SOURCE Coccidioides immitis
ORGANISM Coccidioides immitis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE 1 (bases 1 to 737)
AUTHORS Gardner, M.J. and Kirkland, T.
TITLE Generation of ESTs from Coccidioides immitis spherule cDNA library
JOURNAL Unpublished
COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org.
Location/Qualifiers
1..737
/organism="Coccidioides immitis"
/mol_type="mRNA"
/db_xref="taxon:5501"
/clone="CIAAC67"
/dev_stage="spherule"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 136 a 271 c 168 g 162 t
ORIGIN
Query Match 100.0%; Score 52; DB 10; Length 737;
Best Local Similarity 100.0%; Pred. No. 3.5e-17;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAGTTCTCTCAGCGCTCATCGCTCTGCTGCGTCCGCGCTCCGCACTG 52
Db 99 ATGCAGTTCTCTCAGCGCTCATCGCTCTGCTGCGTCCGCGCTCCGCACTG 150

RESULT 6
BQ499065 311 bp mRNA linear EST 31-OCT-2002
LOCUS EST08290 Pb0001 Paracoccidioides brasiliensis cDNA, mRNA sequence.
DEFINITION BQ499065
ACCESSION BQ499065
VERSION BQ499065.1 GI:24452039
KEYWORDS EST.
SOURCE Paracoccidioides brasiliensis
ORGANISM Paracoccidioides brasiliensis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Oryzenales; mitosporic Oryzenales; Paracoccidioides.
1 (bases 1 to 311)
Goldman,G.H., Marques,E.R., Ribeiro,D.C.D., Bernardes,L.A., Puccia
,R., Travassos,L.R., Nobrega,F.G., Nobrega,M.P., Savoldi-Barbosa,M.
, Semighini,C.P. and Goldman,M.H.
The Paracoccidioides brasiliensis EST genome project
Eukaryotic Cell, (2002) in press
Contact: Gustavo Henrique Goldman
Laboratory of Molecular Biology
Universidade de Sao Paulo - USP - FCFRP
Av do Cafe S/N, CEP. 14040-903, Ribeirao Preto - SP, Brazil
Email: ggoldman@usp.br.
Location/Qualifiers

FEATURES
source
1..311
/organism="Paracoccidioides brasiliensis"
/mol_type="mRNA"
/db_xref="taxon:121759"
/clone_lib="PB0001"
BASE COUNT
73 a 88 c 59 g 91 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY
1 ATGCAGTTCTCTCAGCTCTCAT 23
|||||||
91 ATGCAGTTCTCTCAGCTCTCAT 113

RESULT 7
BP112708/c
-LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BP112708 133 bp mRNA linear EST 11-FEB-2003
BP112708 ORCS1477 5', mRNA sequence.
BP112708
BP112708.1 GI:28314998
EST.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 133)
Ishiwata,H., Katsuna,S., Kizaki,K., Patel,O.V., Nakano,H.,
Takahashi,T., Imai,K., Hirasawa,A., Shiojima,S., Ikawa,H., Suzuki
,Y., Tsujimoto,G., Izaike,Y., Todoroki,J. and Hashizume,K.
Characterization of gene expression profiles in early bovine
pregnancy using a custom cDNA microarray
Mol. Reprod. Dev. 65 (1), 9-18 (2003)
22544902
1258628
Contact: Gozoh Tsujimoto
Department of Molecular, Cell Pharmacology
National Research Institute for Child Health and Development
3-35-31 Tachido, Setagaya, Tokyo 154-8567, Japan
Tel: 81-3-3149-2476
Fax: 81-3-3149-1252
Email: gtsujimoto@nch.go.jp
This work was performed to collaborate with Developmental Biology
Department, National Institute of Agricultural Sciences. Address:
2 Ikenodai, Tsukuba, Ibaraki, 305-8602 Japan. Phone & Fax:
81-29-838-8633 e-mail: kazuh@affrc.go.jp
This work was funded by Organized Research Combination System (ORCS
) Project of Ministry of Education, Culture, Sports, Science and
Technology.
Location/Qualifiers
1..133
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="ORCS13477"
/tissue_type="mixture of uterus and placenta"

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY
4 CAGTTCTCAGCTCTCA 22
|||||||
67 CAGTTCTCAGCTCTCA 49

RESULT 8
AUI63553
-LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AUI63553 314 bp mRNA linear EST 03-APR-2002
AUI63553 Rice panicle at flowering stage Oryza sativa (japonica
cultivar-group) cDNA clone E1586, mRNA sequence.
AUI63553
AUI63553.1 GI:11170958
EST.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Euhartioideae; Oryzaceae; Oryza.
1 (bases 1 to 314)
Sasaki,T. and Yamamoto,K.
Rice cDNA from panicle at flowering stage (2000)
Unpublished
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@br.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT "RGP".
Location/Qualifiers
1..314
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E1586"
/dev_stage="flowering stage"
/clone_lib="Rice panicle at flowering stage"
/note="Organ: panicle; Rice cDNA from panicle at flowering
stage"

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY
24 CGCTCTGCTGCTGCCGC 42
|||||||
187 CGCTCTGCTGCTGCCGC 205

RESULT 9
AUI6766/c
-LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AUI6766 460 bp mRNA linear EST 03-APR-2002
AUI6766 Rice callus (2001) Oryza sativa (japonica cultivar-group)
cDNA clone C50247, mRNA sequence.
AUI6766
AUI6766.1 GI:12405165
EST.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriarthroideae; Oryzaceae; Oryza.
1 (bases 1 to 460)
Sasaki, T. and Yamamoto, K.
Rice cDNA from callus (2001)
Unpublished
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abrr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
PROJECT = "RGP".

FEATURES
source
1..460
Location/Qualifiers

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

1..460
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="C50247"
/issue_type="calus"
/clone_1ib="Rice callus (2001)"
123 c 149 g 76 t 2 others

BASE COUNT
ORIGIN

110 a 123 c 149 g 76 t 2 others

Query Match
Best Local Similarity 100.0%; Score 19; DB 9; Length 460;
Pred. No. 7.5; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db
24 CGCTCTCGTCGTCGCCGC 42
|||||
88 CGCTCTCGTCGTCGCCGC 70
|||||

RESULT 10
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AU031845 468 bp mRNA linear EST 01-APR-2002
AU031845 Rice root Oryza sativa (japonica cultivar-group) cDNA
clone R2118_62, mRNA sequence.
AU031845
AU031845.1 GI:3767818
EST.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriarthroideae; Oryzaceae; Oryza.
1 (bases 1 to 468)
Minobe, Y. and Sasaki, T.
Rice cDNA from root
Unpublished
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abrr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
PROJECT = "RGP".
POLYA=No.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1..468
Location/Qualifiers

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

1..468
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="R2118_62"
/clone_1ib="Rice root"
/note="Prepared from seedling root."

BASE COUNT
ORIGIN

130 a 99 c 118 g 121 t

Query Match
Best Local Similarity 100.0%; Score 19; DB 9; Length 468;
Pred. No. 7.6; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db
14 ACGCTCTCATCGCTCTCGT 32
|||||
101 ACGCTCTCATCGCTCTCGT 83
|||||

RESULT 11
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

A0510461 585 bp DNA linear GSS 04-MAY-1999
nbxb0095B03r CUGI Rice BAC Library Oryza sativa (japonica
cultivar-group) genomic clone nbxb0095B03r, genomic survey
sequence.
A0510461
A0510461.1 GI:4733065
GSS.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriarthroideae; Oryzaceae; Oryza.
1 (bases 1 to 585)
Wing, R. A. and Dean, R. A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished
Contact: Wing RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAACACCATGATGACCATG
Class: BAC ends
High quality sequence stop: 467.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished
Contact: Wing RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAACACCATGATGACCATG
Class: BAC ends
High quality sequence stop: 467.

FEATURES
source
1..585
Location/Qualifiers

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

1..585
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="japonica"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="nbxb0095B03r"
/issue_type="leaf"
/lab_host="E. coli DH10B"
/clone_1ib="CUGI Rice BAC library"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
HindIII; Rice is one of two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 Kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9%. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."

BASE COUNT
ORIGIN

110 a 203 c 150 g 122 t

Query Match
Best Local Similarity 100.0%; Score 19; DB 28; Length 585;
Pred. No. 7.9; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 CTCATCGCTCTCGTCG 37
 |||||
 160 CTCATCGCTCTCGTCG 178

RESULT 12
 CB684405 758 bp mRNA linear EST 09-APR-2003
 LOCUS OSJNEF13003.3 r OSJNEF Oryza sativa (japonica cultivar-group) cDNA

DEFINITION
 CB684405
 OSJNEF13003.3 r OSJNEF Oryza sativa (japonica cultivar-group) cDNA

ACCESSION
 CB684405

VERSION
 CB684405.1 GI:29688130

KEYWORDS
 EST.

SOURCE
 ORGANISM
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 758)
 Jantanasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
 Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.
 Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea
 Unpublished
 Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR Primers
 FORWARD: gta aaa cga cgg cca gtc
 BACKWARD: gga aac agc tat gac cat g
 Plate: 13 row: O column: 03
 Seq primer: gga aac agc tat gac cat g.

FEATURES
 source
 Location/Qualifiers
 1..758
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNEF13003"
 /class_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEF"
 /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
 XhoI; Uninfected Control"

BASE COUNT
 137 a 246 c 198 g 177 t

Query Match 36.5%; Score 19; DB 14; Length 758;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 CGCTCTCGTCGCTCCGCGC 42
 |||||
 406 CGCTCTCGTCGCTCCGCGC 424

RESULT 13
 CB939902 767 bp mRNA linear EST 29-APR-2003
 LOCUS IPCGJx14.9 G08.23 IPCGJx14 Ictalurus punctatus cDNA clone

DEFINITION
 IPCGJx14.9 G08.23 IPCGJx14 Ictalurus punctatus cDNA clone

ACCESSION
 CB939902

VERSION
 CB939902.1 GI:30225293

KEYWORDS
 EST.

SOURCE
 ORGANISM
 Ictalurus punctatus (channel catfish)
 Ictalurus punctatus

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Bengten,E., Khayat,M., Middleton,D., Waldbieser,G., Askovic,S.,
 Jensen,K.T., Warr,G., Miller,N., Cizm,L.M. and Wilson,M.
 Identification of expressed genes in mixed leukocyte culture and in
 a macrophage cell line in channel catfish, Ictalurus punctatus
 Unpublished
 Contact: Waldbieser GC
 Catfish Genetics Research Unit
 USDA-Agricultural Research Service
 141 Experiment Station Road, Stoneville, MS 38776, USA
 Tel: 662 686 3593
 Fax: 662 686 3567
 Email: gwaldbieser@ars.usda.gov
 Single pass sequencing. Bases called with Phred v0.000925.c. Low
 quality bases and vector trimmed with Lucy v1.16.
 Plate: 9 row: G column: 8
 Seq primer: T7.

FEATURES
 source
 Location/Qualifiers
 1..767
 /organism="Ictalurus punctatus"
 /mol_type="mRNA"
 /strain="commercial"
 /db_xref="taxon:7998"
 /clone="IPCGJx14.9 G08"
 /cell_type="macrophage"
 /cell_line="42TA"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="IPCGJx14"
 /note="Vector: pSport1; Site_1: Sal I; Site_2: Not I;
 Primary library"

BASE COUNT
 250 a 169 c 271 g 77 t

Query Match 36.5%; Score 19; DB 14; Length 767;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GCTCTCATCGCTCTCGTCG 34
 |||||
 506 GCTCTCATCGCTCTCGTCG 488

RESULT 14
 BE039817 978 bp mRNA linear EST 07-JUN-2000
 LOCUS OC08E11 OC Oryza sativa cDNA 5', mRNA sequence.

DEFINITION
 BE039817

ACCESSION
 BE039817

VERSION
 BE039817.1 GI:8334833

KEYWORDS
 EST.

SOURCE
 ORGANISM
 Oryza sativa
 Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 978)
 Bohmert,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferreira
 H., Kawasaki,S., McCollough,A., Michalowski,C.B., Palacio,C.,
 Scara,G., Wheeler,M. and Zepeda,G.R.
 Functional Genomics of Plant Stress Tolerance
 Unpublished
 Contact: Michalowski,C.B.
 University of Arizona
 Bio Sciences West room 513, Tucson, AZ 85721, USA
 Tel: 520-621-7982
 Fax: 520-621-1697
 Email: cbm@u.arizona.edu
 Best blastx match: 'emb|CAA1636.1| (A021633) predicted protein
 [Arabidopsis thaliana] 51 1e-05'. An open reading frame exists.
 Location/Qualifiers

FEATURES

source

1. .978
/organism="Oryza sativa"
/mol_type="mRNA"
/strain="pokkali"
/db_xref="taxon:4530"
/tissue_type="roots"
/dev_stage="1 week"
/clone_1id="OC"
/note="no stress"
BASE COUNT 232 a 188 c 313 g 244 t 1 others
ORIGIN

Query Match 36.5%; Score 19; DB 10; Length 978;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 ACCTCTCATCGCTCTCTCT 32
|||||
Db 168 ACCTCTCATCGCTCTCTCT 150
|||||

RESULT 15

LOCUS BG163853 1712 bp mRNA linear EST 06-FEB-2001
DEFINITION 602343204F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4453297 5',
mRNA sequence.
ACCESSION BG163853
VERSION BG163853.1 GI:12670556
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1712)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
AUTHORS Contact: Robert Strausberg, Ph.D.
TITLE Email: cgabs-remail.nih.gov
COMMENT Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov

Plate: LLM10243 row: h column: 02
High quality sequence stop: 5.
Location/Qualifiers

FEATURES

source

1. .1712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_1id="IMAGE:4453297"
/tissue_type="hypertrophoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1id="NIH_MGC_89"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; Oligo-dT primed.
Average insert size 1.3 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 691 a 415 c 350 g 256 t
ORIGIN

Query Match 36.5%; Score 19; DB 10; Length 1712;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 CGCTCTCATCGCTCTCTCT 33
|||||
Db 1632 CGCTCTCATCGCTCTCTCT 1614
|||||

RESULT 16

LOCUS A0845447/c 368 bp DNA linear GSS 25-MAY-2001
DEFINITION LMAJFV1_lm25f07.y1 Leishmania major FV1 random genomic library
Leishmania major genomic clone LMAJFV1_lm25f07 5', genomic survey
sequence.
ACCESSION A0845447
VERSION A0845447
KEYWORDS GSS.
SOURCE Leishmania major
ORGANISM Leishmania major
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;

REFERENCE 1 (bases 1 to 368)
Akopyants, N.S., Clifton, S.W., Martin, J., Pape, D., Wylie, T., Li, L.,
Kissinger, J.C., Rose, D.S., and Beverley, S.M.
A survey of the Leishmania major Friedlin strain V1 genome by
shotgun sequencing: a resource for DNA microarrays and expression
profiling

Mol. Biochem. Parasitol. 113 (2), 337-340 (2001)

JOURNAL MEDLINE PUBMED

Other GSSs: lm25f07.x1
Contact: Akopyants, NS / Beverley, SM
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: eatw@wustl.edu

Library construction: Natalia S. Akopyants, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
If using this information please cite:
N.S. Akopyants and S.M. Beverley 'A survey of the Leishmania major
Friedlin strain V1 genome by shotgun sequencing' and the Washington
University Genome Sequencing Center for information on obtaining
clone material please contact: Natalia S. Akopyants Ph.D.
(natalia@borcim.wustl.edu) and/or Stephen M. Beverley Ph.D.
(beverley@borcim.wustl.edu)
Seq primer: -40RP from Gibco
Class: shotgun
High quality sequence stop: 358.
Location/Qualifiers

FEATURES

source

1. .368
/organism="Leishmania major"
/mol_type="genomic DNA"
/strain="Friedlin strain V1"
/db_xref="taxon:5664"
/clone_1id="LMAJFV1_lm25f07"
/lab_host="TOP10 (Invitrogen)"
/clone_1id="Leishmania major FV1 random genomic library"
/note="Vector: pZero-2 (Invitrogen); Site 1: EcoRV;
Genomic DNA was isolated from stationary phase cells. For
this library, DNA was sheared to give a tight size
distribution of 1-1.5kb fragments, blunt-ended with T4 DNA
polymerase, dephosphorylated with Shrimp Alkaline
Phosphatase and ligated into pZero-2 vector's EcoRV site."

BASE COUNT 90 a 94 c 123 g 61 t
ORIGIN

Query Match 34.6%; Score 18; DB 28; Length 368;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 TCATCGCTCTCGCTCGCTG 37
|||||
Db 135 TCATCGCTCTCGCTCGCTG 118
|||||

RESULT 17

LOCUS B1568671/c 546 bp mRNA linear EST 06-SEP-2001
B1568671


```

DEFINITION RH40331.5prime RH Drosophila melanogaster normalized Head pflc-1
ACCESSION Drosophila melanogaster cDNA clone RH40331 5, mRNA sequence.
VERSION B1568671
KEYWORDS B1568671.1 GI:15460093
SOURCE EST
ORGANISM Drosophila melanogaster (fruit fly)
REFERENCE Drosophila melanogaster
AUTHORS Eukaryota: Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephyroidae; Drosophilidae; Drosophila.
1 (bases 1 to 546)
Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
J., Champagne,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Mista,S.,
Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Phonenuavong,S., Wan,K., Yu,C., Lewis,S.E., Celnikier,S. and Rubin
G.M.
BDGP/HHMI RH Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003466: arm:2R [19937589,20220687]
estimated-cyto:60E8-60F4: 08/18/2001
Plate: RH.403 row: C column: 7
High quality sequence stop: 508.
Location/Qualifiers
1..546
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RH40331"
/sex="male and female"
/dev_stage="Adult"
/lab_host="DHS-alpha Tona"
/clone_lib="RH Drosophila melanogaster normalized Head
pflc-1"
/notes="Organ: head; Vector: pflc1; Site 1: XhoI; Site 2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT 100 a 148 c 176 g 120 t 2 others
ORIGIN
Query Match 34.6%; Score 18; DB 12; Length 546;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 CATGCTCTCGTCGCTGC 38
|||||
Db 231 CATGCTCTCGTCGCTGC 214

RESULT 18
B1626323/c 558 bp mRNA linear EST 06-SEP-2001
LOCUS RH22614.5prime RH Drosophila melanogaster normalized Head pflc-1
DEFINITION Drosophila melanogaster cDNA clone RH22614 5, mRNA sequence.
ACCESSION B1583892
VERSION B1583892
KEYWORDS B1583892.1 GI:15475314
SOURCE EST
ORGANISM Drosophila melanogaster (fruit fly)
REFERENCE Drosophila melanogaster
AUTHORS Eukaryota: Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephyroidae; Drosophilidae; Drosophila.
1 (bases 1 to 558)
Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
J., Champagne,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Mista,S.,
Phonenuavong,S., Wan,K., Yu,C., Lewis,S.E., Celnikier,S. and Rubin
G.M.
BDGP/HHMI RH Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003786: arm:2R [607836,924593]
estimated-cyto:41F5-41F9: 08/17/2001
Plate: RH.226 row: B column: 2
High quality sequence stop: 501.
Location/Qualifiers
1..558
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RH22614"
/sex="male and female"
/dev_stage="Adult"
/lab_host="DHS-alpha Tona"
/clone_lib="RH Drosophila melanogaster normalized Head
pflc-1"
/notes="Organ: head; Vector: pflc1; Site 1: XhoI; Site 2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT 101 a 152 c 180 g 124 t 1 others
ORIGIN
Query Match 34.6%; Score 18; DB 12; Length 558;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 CATGCTCTCGTCGCTGC 38
|||||
Db 231 CATGCTCTCGTCGCTGC 214

RESULT 19
B1626323/c 610 bp mRNA linear EST 07-SEP-2001
LOCUS RH67006.5prime RH Drosophila melanogaster normalized Head pflc-1
DEFINITION Drosophila melanogaster cDNA clone RH67006 5, mRNA sequence.
ACCESSION B1626323
VERSION B1626323
KEYWORDS B1626323.1 GI:15521848
SOURCE EST
ORGANISM Drosophila melanogaster (fruit fly)
REFERENCE Drosophila melanogaster
AUTHORS Eukaryota: Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephyroidae; Drosophilidae; Drosophila.
1 (bases 1 to 610)
Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
J., Champagne,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Mista,S.,
Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Phonenuavong,S., Wan,K., Yu,C., Lewis,S.E., Celnikier,S. and Rubin
G.M.
BDGP/HHMI RH Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003786: arm:2R [607836,924593]
estimated-cyto:41F5-41F9: 08/23/2001
Plate: RH.670 row: A column: 6

```

```

Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Phonenuavong,S., Wan,K., Yu,C., Lewis,S.E., Celnikier,S. and Rubin
G.M.
BDGP/HHMI RH Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003786: arm:2R [607836,924593]
estimated-cyto:41F5-41F9: 08/17/2001
Plate: RH.226 row: B column: 2
High quality sequence stop: 501.
Location/Qualifiers
1..558
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RH22614"
/sex="male and female"
/dev_stage="Adult"
/lab_host="DHS-alpha Tona"
/clone_lib="RH Drosophila melanogaster normalized Head
pflc-1"
/notes="Organ: head; Vector: pflc1; Site 1: XhoI; Site 2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT 101 a 152 c 180 g 124 t 1 others
ORIGIN
Query Match 34.6%; Score 18; DB 12; Length 558;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 CATGCTCTCGTCGCTGC 38
|||||
Db 231 CATGCTCTCGTCGCTGC 214

RESULT 19
B1626323/c 610 bp mRNA linear EST 07-SEP-2001
LOCUS RH67006.5prime RH Drosophila melanogaster normalized Head pflc-1
DEFINITION Drosophila melanogaster cDNA clone RH67006 5, mRNA sequence.
ACCESSION B1626323
VERSION B1626323
KEYWORDS B1626323.1 GI:15521848
SOURCE EST
ORGANISM Drosophila melanogaster (fruit fly)
REFERENCE Drosophila melanogaster
AUTHORS Eukaryota: Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephyroidae; Drosophilidae; Drosophila.
1 (bases 1 to 610)
Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
J., Champagne,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Mista,S.,
Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Phonenuavong,S., Wan,K., Yu,C., Lewis,S.E., Celnikier,S. and Rubin
G.M.
BDGP/HHMI RH Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003786: arm:2R [607836,924593]
estimated-cyto:41F5-41F9: 08/23/2001
Plate: RH.670 row: A column: 6

```

FEATURES
source High quality sequence stop: 525.
Location/Qualifiers
1. .610
/organism="Drosophila melanogaster"
/mol_type="mrna"
/db_xref="taxon:7227"
/clone="RH67006"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5-alpha Tona"
/clone_1lb="RH Drosophila melanogaster normalized Head
p1c-1"
/note="Organ: head; Vector: p1c1; Site: 1: XhoI; Site 2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."

BASE COUNT 115 a 163 c 198 g 133 t 1 others

ORIGIN

Query Match 34.6%; Score 18; DB 12; Length 610;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 CATCGCTCTCGTCGCTGC 38
|||||
231 CATCGCTCTCGTCGCTGC 214
|||||

RESULT 20
LOCUS A2208101 936 bp DNA linear GSS J1-AUG-2000
DEFINITION SP_0136_B1_F01_T7A Strongylocentrotus purpuratus, purple sea urchin
'Sperm genomic BAC library Strongylocentrotus purpuratus genomic
clone Plate=136 Col=1 Row=L, genomic survey sequence.

ACCESSION A2208101 GI:8420375
KEYWORDS GSS.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Echinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 936)
Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,
Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray,
G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J., Davidson,E.H. and
Hood,L.
A sea urchin genome project: Sequence scan, virtual map, and
additional resources
Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)

JOURNAL MEDLINE 20402566
PUBMED 10920195
CONTACT: Cameron, RA, Davidson, EH, Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 136 row: L column: 1
Seq primer: T7
Classes: BAC ends
High quality sequence stop: 936.
Location/Qualifiers
1. .936
/organism="Strongylocentrotus purpuratus"
/mol_type="genomic DNA"
/db_xref="taxon:7668"
/clone="Plate=136 Col=1 Row=L"
/clone_1lb="Strongylocentrotus purpuratus, purple sea
urchin, sperm genomic BAC library"
/note="Organ: sperm; Vector: BAC3.6; BAC Clones in E-Coli
DH10B"

FEATURES
source

BASE COUNT 227 a 231 c 163 g 315 t

ORIGIN

Query Match 34.6%; Score 18; DB 28; Length 936;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 TCTCATCGCTCTCGTCGC 35
|||||
Db 607 TCTCATCGCTCTCGTCGC 624
|||||

RESULT 21
LOCUS CNS01736 1012 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BACN17C24 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL107532
VERSION AL107532.1 GI:5627836
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1012)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBeloBAC11.

FEATURES
source
Location/Qualifiers
1. .1012
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN17C24"
/clone_1lb="DrosBAC"
/plasmid="pBeloBAC11"
/note="end : T7"

BASE COUNT 192 a 287 c 249 g 233 t 51 others

ORIGIN

Query Match 34.6%; Score 18; DB 29; Length 1012;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 CATCGCTCTCGTCGCTGC 38
|||||
Db 953 CATCGCTCTCGTCGCTGC 970
|||||

RESULT 22
LOCUS BF307868 1083 bp mRNA linear EST 21-NOV-2000
DEFINITION 601890683F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4131836 5',
mRNA sequence.

ACCESSION BF307868
VERSION BF307868.1 GI:11255033
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1083)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>,
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
 Plate: LNCM1030 row: e column: 21
 High quality sequence stop: 466.
 Location/Qualifiers
 1..1083
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4131836"
 /issue_type="rhabdomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH MGC 17"
 /note="Organ: muscle; Vector: pORF7; Site 1: EcoRI;
 Site 2: XhoI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 218 a 363 c 318 g 184 t
 ORIGIN
 Query Match 34.6%; Score 18; DB 10; Length 1083;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 TGCAGTTCTCTCAGCGTC 19
 |||||
 414 TGCAGTTCTCTCAGCGTC 431
 RESULT 23
 CNS00ATX 1101 bp DNA linear GSS 04-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #
 DEFINITION BACR22L05 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL056234.1 GI:4936801
 VERSION AL056234.1
 SOURCE GSS.
 ORGANISM Drosophila melanogaster (fruit fly)
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)
 Genoscope.
 Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osoegawa and
 Aaron Mamooser in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain Y2; cn bw sp. the same strain used for the BDGP's
 p1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
 Location/Qualifiers
 1..1101
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR22L05"
 /clone_1lb="RPCI-98"
 /note="end : T7"
 BASE COUNT 254 a 295 c 253 g 226 t 73 others
 ORIGIN
 Query Match 34.6%; Score 18; DB 29; Length 1101;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 CATCGCTCTGTCGTCGTC 38
 |||||
 Db 608 CATCGCTCTGTCGTCGTC 625
 RESULT 24
 CNS014JW 1101 bp DNA linear GSS 26-JUL-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
 DEFINITION BACN12F17 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL104246.1 GI:5615857
 VERSION AL104246.1
 SOURCE GSS.
 ORGANISM Drosophila melanogaster (fruit fly)
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)
 Genoscope.
 Direct Submission
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billard at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBelobAC11.
 Location/Qualifiers
 1..1101
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACN12F17"
 /clone_1lb="DrosBAC"
 /plasmid="pBelobAC11"
 /note="end : T7"
 BASE COUNT 236 a 261 c 315 g 255 t 34 others
 ORIGIN
 Query Match 34.6%; Score 18; DB 29; Length 1101;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 CATCGCTCTGTCGTCGTC 38
 |||||
 Db 175 CATCGCTCTGTCGTCGTC 158

RESULT 25
LOCUS B1123995 343 bp mRNA linear EST 31-DEC-2001
DEFINITION 1033P32P Populus leaf cDNA library Populus tremula x Populus tremuloides cDNA, mRNA sequence.
ACCESSION B1123995
VERSION B1123995.1 GI:18007970
KEYWORDS EST.
SOURCE Populus tremula x Populus tremuloides
ORGANISM Populus tremula x Populus tremuloides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid 1; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 343)
Hertzberg, M., Aspeborg, H., Erlandsson, R., Bjorkbacka, H., Hilonen, T., Karlsson, J., Teeri, T., Gustafsson, P., Bahlert, R., Jansson, S., Nilsson, O., Sundberg, B., Nilsson, P., Uhlen, M., Sandberg, G. and Lundberg, J.
Gene expression in Populus
Unpublished
Contact: Erlandsson R
Department of Biotechnology
Royal Institute of Technology
Teknikringen 30, Stockholm S-10044, Sweden
Tel: 46 8 790 8287
Fax: 46 8 245452
Email: riket@biochem.kth.se.

FEATURES
source
1..343
/organism="Populus tremula x Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:47664"
/clone_lib="Populus leaf cDNA library"
/note="Organ: leaf"

BASE COUNT 75 a 89 c 58 g 121 t
ORIGIN

Query Match 32.7% Score 17; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GCTCTCATCGCTCTCGT 32
|||||
6 GCTCTCATCGCTCTCGT 22
|||||

Db 6 GCTCTCATCGCTCTCGT 22

RESULT 26
LOCUS CD262900/c 359 bp mRNA linear EST 23-MAY-2003
DEFINITION pSMA019XK10f.192458 sMA: Phytophthora sojae grown in synthetic medium
Phytophthora sojae cDNA clone sMA019K10 5, mRNA sequence.
ACCESSION CD262900
VERSION CD262900.1 GI:31050727
KEYWORDS EST.
SOURCE Phytophthora sojae
ORGANISM Phytophthora sojae
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
Phytophthora
Phytophthora
1 (bases 1 to 359)
Tyler, B.M., Judelson, H.S., Gijzen, M., Dean, R.A. and Waugh, M.E.
USDA-IFAFS: Expression of Phytophthora sojae genes during infection
Unpublished
Contact: Tyler B
Tyler lab
VBI
1890 Pratt Dr., Blacksburg, VA 24061, USA
Tel: 540-231-7318
Email: bmtylet@vt.edu
PCR Primers
FORWARD: BK reverse
Plate: 019 row: K column: 10

Seq primer: BK reverse
High quality sequence scop: 359.
Location/Qualifiers
1..359
/organism="Phytophthora sojae"
/mol_type="mRNA"
/db_xref="taxon:67593"
/clone="sMA019K10"
/issue_type="mycelium"
/cell_line="P6497"
/dev_stage="mycelium"
/clone_lib="sMA: Phytophthora sojae grown in synthetic medium"
/note="vector: pBK-CMW; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 73 a 125 c 106 g 55 t
ORIGIN

Query Match 32.7% Score 17; DB 14; Length 359;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 GCTCTGTCGTCGCCG 41
|||||
Db 128 GCTCTGTCGTCGCCG 112
|||||

RESULT 27
LOCUS BY076213/c 374 bp mRNA linear EST 06-DEC-2002
DEFINITION BY076213 RIKEN full-length enriched, pooled tissues, adult spleen, etc. Mus musculus cDNA clone K63003015 5', mRNA sequence.
ACCESSION BY076213
VERSION BY076213.1 GI:26177732
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 374)
Oikarai, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bul, C., Hume, D.A., Quackenbush, J., Schirni, L.M., Kanapin, A., Matsuda, H., Batalov, S., Belser, K.W., Blake, J.A., Brad, D., Brusic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Paven, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C.A., Setu, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Whitley, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wysshaw-Borle, A., Yanoisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashitume, M., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

TITLE
JOURNAL MEDLINE
COMMENT PUBMED
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

```

/dev_stage="larva"
/clone_1lb="Amblyomma americanum larva lambda Zap Express"
/notes="Vector: PBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      92 a      91 c      118 g      101 t      4 others
ORIGIN

Query Match      32.7%; Score 17; DB 10; Length 406;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCACTCTCTCAGCC 17
|||||
154 ATGCACTCTCTCAGCC 138

RESULT 30
AM465276/c      425 bp      mRNA      linear      EST 24-FEB-2000
DEFINITION      BP230018B10F10 Soares normalized bovine placenta Bos taurus CDNA
ACCESSION      AM465276
VERSION      AM465276
KEYWORDS      EST
SOURCE      Bos taurus (cow)
ORGANISM      Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 425)
Lewin,H.A., Soares,M.B., Rebeiz,M., Pardinas,J., Liu,L. and Larson
J.H.
Bovine ESTs
Unpublished
Contact: Lewin, H. A.
W. M. Keck Center for Comparative and Functional Genomics
University of Illinois at Urbana-Champaign
340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL
61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Funding for cattle EST sequencing was provided by the USDA National
Research Initiative, Animal Genome Resource Grant AG 99-3205-8534
to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED
from Washington University Genome Center. Vector Trim: 9;
Cross match from Washington University Genome Center PHRAP suite.
Sequences submitted are vector free and at least 200 bp in length.
PCR Primers
FORWARD: TATACGACTCACTATAGG
BACKWARD: ATTACCTCTACTAAG
Insert Length: 425 Std Error: 0.00
Plate: BP230018B10 row: F column: 10
Seq primer: AGCGATACCAATTTCACACAGGA
High quality sequence stop: 425.
Location/Qualifiers
source
1. 425
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="BP230018B10F10"
/sex="female"
/lab_host="DH10B"
/clone_1lb="Soares normalized bovine placenta"
/notes="Organ: Placenta; Vector: pTZ193pac; Site_1: EcoRI;
Site_2: NotI; The CDNA library was contributed by the
Soares laboratory and it was constructed and normalized
as described by Bonaldo, M.F., Lennon, G. and Soares,
M.B. (1996), Genome Research 6(9): 791-806. "
BASE COUNT      93 a      116 c      149 g      67 t
ORIGIN

Query Match      32.7%; Score 17; DB 9; Length 425;
Best Local Similarity 100.0%; Pred. No. 84;

```

```

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      22 ATGCACTCTCTCAGCC 38
|||||
335 ATGCACTCTCTCAGCC 319

RESULT 31
AM056821
LOCUS      ST56F10 Pine Triplex shoot tip library Pinus taeda CDNA clone
DEFINITION      ST56F10, mRNA sequence.
ACCESSION      AM056821
VERSION      AM056821.1 GI:5929516
KEYWORDS      EST
SOURCE      Pinus taeda (loblolly pine)
ORGANISM      Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferophyta; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 446)
Whetten,R.W., Kinlaw,C.S., Retzel,E. and Sederoff,R.R.
The Pine Gene Discovery Project
Unpublished
Contact: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh
NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhetten@unity.ncsu.edu
Shoot tip library plates ST01-ST50 were prepared and sequenced at
NCsu. Plates ST51-ST99 were prepared and sequenced at IFG/CSUH.
Informatics was done at the Computational Biology Centers (CBC) at
University of Minnesota.
Seq primer: 5' lambda Triplex2 sequencing primer.
Location/Qualifiers
source
1. 446
/organism="Pinus taeda"
/mol_type="mRNA"
/db_xref="taxon:3352"
/clone="ST56F10"
/lab_host="E. coli BM25.8"
/clone_1lb="Pine Triplex shoot tip library"
/notes="Organ: shoot tips; Vector: lambda Triplex; Site_1:
SfiI (A); Site_2: SfiI (B); Shoot tips (approx. 2 cm from
apex) were collected during the spring, frozen and used
for mRNA isolation. The SMART-PCR method (Clontech) was
used to prepare a library from 1 ug total RNA, using the
lambda Triplex vector. Plasmid subclones in pTriplex were
recovered by cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end."
BASE COUNT      107 a      137 c      79 g      110 t      13 others
ORIGIN

Query Match      32.7%; Score 17; DB 9; Length 446;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      31 GTCGCTGCGCGCCTCGC 47
|||||
80 GTCGCTGCGCGCCTCGC 96

RESULT 32
AM376507
LOCUS      IL3-CT0213-161299-038-F12 CT0213 Homo sapiens CDNA, mRNA sequence.
DEFINITION      AM376507
ACCESSION      AM376507.1 GI:6881168
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 470)
 TITLE HCGP <http://www.ludwig.org.br/ORESTES>.
 JOURNAL The FAPESP/LICR Human Cancer Genome Project
 COMMENT Unpublished
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=113&ct=113-CT0213-161299-038-F12&ct=1999-12-16&ct=1>)
 Seq primer: puc 18 forward
 High quality sequence stop: 383.

FEATURES
 source
 1..470
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_11b="CT0213"
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 116 a 119 c 152 g 83 t
 ORIGIN

Query Match 32.7%; Score 17; DB 9; Length 470;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 ATCGCTCTCGCTCGCTC 38
 |||||
 276 ATCGCTCTCGCTCGCTC 260

RESULT 33
 BQ170403 490 bp mRNA linear EST 26-APR-2002
 LOCUS BQ170403
 DEFINITION WHE1770_E05 J102T Wheat pre-anthesis spike cDNA library Triticum aestivum
 VERSION BQ170403.1 GI:20332226
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticaceae; Triticum.
 1 (bases 1 to 490)
 Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and Tong,J.C.
 The structure and function of the expressed portion of the wheat genomes - pre-anthesis spike cDNA library
 Unpublished
 Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105555773
 Fax: 5105555818
 Email: anderson@pw.usda.gov

TITLE
 JOURNAL
 COMMENT

REFERENCE
 AUTHORS
 JOURNAL
 COMMENT

FEATURES
 source
 1..490
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultiivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone="WHE1770_E05 J10"
 /cissue_type="Spike_before anthesis"
 /dev_stage="Adult plant"
 /lab_host="E. coli S04R"
 /clone_11b="wheat pre-anthesis spike cDNA library"
 /note="Vector: Lambda Uni-ZAP XR, excised phagmid; Site 1: EcoRI; Site 2: XhoI; Plants were grown in the greenhouse. Whole spike with awns trimmed. White, green and yellow anther were collected and total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagmids in the T7 close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 124 a 138 c 117 g 111 t
 ORIGIN

Query Match 32.7%; Score 17; DB 13; Length 490;
 Best Local Similarity 100.0%; Pred. No. 86;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26 CTCTCGTCTGCTCGCGC 42
 |||||
 345 CTCTCGTCTGCTCGCGC 361

RESULT 34
 BE484921/c 505 bp mRNA linear EST 27-MAR-2003
 LOCUS BE484921
 DEFINITION 171741 BARC SBOV Bos taurus cDNA 5', mRNA sequence.
 VERSION BE484921
 KEYWORDS BE484921.1 GI:9604454
 EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
 1 (bases 1 to 505)
 Sonstegard,T., Capuco,A.V., White,J., Van Tassel,J.C.P., Connor,E.E., Cho,J., Sultana,R., Shade,L., Wray,J.E., Wells,K.D. and Quackenbush,J.
 Analysis of bovine mammary gland EST and functional annotation of the Bos taurus gene index
 Mamm. Genome 13 (7), 373-379 (2002)
 22135956
 12140684

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

USDA, ARS, Beltsville Agricultural Research Center
 Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
 Tel: 301 504 8416
 Fax: 301 504 8414
 Email: tads@lps1.barc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -mnscore 18 and -mismatch 12 options.
 PCR PRIMERS
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCACGACGACGACG
 Place: 135 row: J column: 12
 Seq primer: ATTAGTGACACTATAG.

FEATURES

Location/Qualifiers


```

source
1. .505
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/feature_type="pooled"
/lab_host="DH10B"
/clone_id="BARC 5BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."
BASE COUNT      118 a      135 c      175 g      77 t
ORIGIN
Query Match      32.7%; Score 17; DB 10; length 505;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      22 ATCGCTCTCGCGCTGC 38
|||||
db      342 ATCGCTCTCGCGCTGC 326

```

RESULT 35	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
B0821431/c	B0821431	10300932cdp.y1 C. reinhardtii CC-1690, DeFlagellation (normalized), Lambda 2c4p II Chlamydomonas reinhardtii cDNA, mRNA sequence.	B0821431	B0821431		EST.	Chlamydomonas reinhardtii	Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D. Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1030	unpublished	Duke University Durham, NC 27708-1000	Tel: 919 613 8159 Fax: 919 613 8177 Email: chausser@duke.edu.	

FEATURES	source
location/Qualifiers	
1. 506	
/organism="Chlamydomonas reinhardtii"	
/mol_type="mRNA"	
/strain="CC-1690 wild type mt+ 21gr"	
/db_xref="taxon:1055"	
/clone_lib="C. reinhardtii CC-1690, Deflagellation (normalized), lambda Zap II"	
/note="Vector: pluscript II SK-, Site 1: EcoRI, Site 2: XhoI; Deflagellation library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells which had been re-synthesizing flagella for 15, 30 and 60 min after being deflagellated by pH shock. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. Pluscript II SK- plasmids were excised from the lambda Zap clones by superinfection with Exsist (Stratagene) phase. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."	
BASE COUNT	
ORIGIN	
98 a	153 c 168 g 87 t
Query Match	32.7%; Score 17; DB 13; Length 506;

Best Local Similarity	100.0%	Pred No. 87;
Matches	17; Conservative	0; Mismatches
		0; Indels
		0; Gaps
		0;
QY	28 CTCGTCGCTGCCGCCT 44 	
Db	362 CTCGTCGCTGCCGCCT 346	
RESULT 36		
A1988643/c		
LOCUS	A1988643	529 bp mRNA linear EST 30-NOV-2001
DEFINITION	ad5sh05.y1 Gm-cl020 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:	
ACCESSION	Gm-cl020-850 5', mRNA sequence.	
VERSION	A1988643	
KEYWORDS	A1988643.1 GI:5820437	
SOURCE	EST.	
ORGANISM	Glycine max (soybean)	
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosid1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine. 1 (bases 1 to 529) Shoemaker,R., Keim,P., Vodkin,L., Erpelnding,J., Corvelli,V., Khanna A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers Y., Pearson,B., Swailer,T., Gibbons,M., Pape,D., Harvey,N., Schurk R., Ritter,E., Korn,S., Shun,T., Jackson,Y., Cardenas,M., McCann R., Waterson,R. and Wilson,R. Public Soybean EST Project Unpublished Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@waterson.wustl.edu This clone is available through ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntleville, AL 35801 For further information call: (800)-533-4363 or contact via email: cu@resgen.com Insert Length: 810 Std Error: 0.00 Seq primer: -40RP from Gibco High quality sequence stop: 429. Location/Qualifiers 1..529 /organism="Glycine max" /mol_type="mRNA" /db_xref="taxon:3847" /clone="GENOME SYSTEMS CLONE ID: Gm-cl020-850"	

```

//tissue_type="root nodules of greenhouse grown plants"
//lib_host="Xl10-Gold"
//clone_lib="Gm-c1020"
//notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from nodules on the roots of 2.5 month-old Glycine max
'Williams' plants that were greenhouse grown. The cDNA
library was prepared using the Stratagene pBluescript II
SK(+) library construction kit. First strand synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An 'anchor'
nucleotide (V=A, C, or G) was added to the 3' end of the
primer (GAGCAGAGACGAGACGACTGCTCGAG(T)18V) to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in with
cloned Pfu DNA polymerase and size-fractionated with a
400 bp cutoff, using a Siseasp 400 Spin column from
Pharmacia. The column eluent was ligated to EcoRI adaptors
and phosphorylated. The XhoI sites in the cDNA would be
protected by their hemimethylated status. The cDNA
constructs were size-fractionated with a 500 bp cutoff,
using GibcoBRL Life Technologies' cDNA size predigested

```


vector (pBluescript II SK-) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene. Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=56) and sequence (n=16). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT 94 a 157 c 179 g 98 t 1 others

Query Match 32.7%; Score 17; DB 9; Length 529;

Best Local Similarity 100.0%; Pred. No. 88;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 TCTCATGCTCTGTCG 34
|||||
529 TCTCATGCTCTGTCG 513

RESULT 37
AV668034 542 bp mRNA linear EST 28-NOV-2001
LOCUS AV668034 Bos taurus ovary fetus Bos taurus cDNA clone EIOV019D09
DEFINITION 5', mRNA sequence.

ACCESSION AV668034 GI:99327779
VERSION AV668034.1
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (bases 1 to 542)
Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H. and Sugimoto,Y.

REFERENCE
AUTHORS
TITLE Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs

JOURNAL Nucleic Acids Res. 29 (22), E108 (2001)

MEDLINE 21570554

PUBMED 11713328

COMMENT Contact: Yoshikazu Sugimoto

Animal Genetics Division

Shirakawa Institute of Animal Genetics

Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan

Tel: 81-248-25-5641

Fax: 81-248-25-5725

Email: kazusugi@cocoa.ocn.ne.jp

Single pass sequencing.

This clone was obtained from a polyA-deleted cDNA library.

FEATURES

source

Location/Qualifiers

1..542

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/clone="EIOV019D09"

/tissue_type="ovary"

/dev_stage="fetus"

/lab_host="DH10B"

/clone_lib="Bos taurus ovary fetus"

/note="Vector: pZL1, Site_1: SalI, Site_2: NotI, Poly A

was deleted from a NotI site"

BASE COUNT 121 a 153 c 187 g 81 t.

ORIGIN

Query Match 32.7%; Score 17; DB 9; Length 542;

Best Local Similarity 100.0%; Pred. No. 88;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 ATGCTCTCTGCTGCTG 38
|||||

Db 375 ATGCTCTCTGCTGCTG 359

RESULT 38
A1542489/c
LOCUS A1542489

544 bp mRNA linear EST 23-APR-2001
SD08831.5prime SD Drosophila melanogaster Schneider L2 cell culture

DEFINITION

POT2 Drosophila melanogaster cDNA clone SD08831.5 similar to

AJ010298: Drosophila melanogaster retrotransposon-like element,

mRNA sequence.

ACCESSION A1542489

VERSION A1542489.2 GI:13771703

KEYWORDS EST.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 544)

REFERENCE Harvey,D., Broksstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,

Lewis,S. and Rubin,G.M.

BDGP/HMT Drosophila EST Project

Unpublished

On Mar 19, 1999 this sequence version replaced gi:4459862.

Other ESTs: SD08831.3prime

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu

hit genomic AB003350: arm:U (4635408,4651408) estimated-cyto.?

04/13/2001

Plate: SD.68 row: C column: 7

High quality sequence stop: 459

POLYA=NO.

FEATURES

source

Location/Qualifiers

1..544

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="SD08831"

/lab_host="DH5-alpha"

/clone_lib="SD Drosophila melanogaster Schneider L2 cell

culture POT2"

/note="Vector: POT2, Site_1: EcoRI, Site_2: XhoI, sized

fractionated cDNAs were directly ligated into POT2.

Plasmid cDNA library."

BASE COUNT 119 a 148 c 171 g 106 t

ORIGIN

Query Match 32.7%; Score 17; DB 9; Length 544;

Best Local Similarity 100.0%; Pred. No. 88;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 CATGCTCTGCTGCTG 37
|||||

Db 471 CATGCTCTGCTGCTG 455

RESULT 39
AV597173/c
LOCUS AV597173

546 bp mRNA linear EST 27-NOV-2001
AV597173 Bos taurus cartilage fetus Bos taurus cDNA clone

DEFINITION

EICA030D01.5', mRNA sequence.

ACCESSION AV597173

VERSION AV597173.1 GI:9714516

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 546)

REFERENCE Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.

and Sugimoto,Y.

TITLE Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs

JOURNAL Nucleic Acids Res. 29 (22), E108 (2001)

MEDLINE 21570554

PUBMED 11713328

COMMENT Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5725
Fax: 81-248-25-5725
Email: kazusugi@cocoa.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.

FEATURES

source

1..546
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1CA030D01"
/tissue_type="cartilage"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_11b="Bos taurus cartilage fetus"
/note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A was deleted from a NotI site"

BASE COUNT 126 a 148 c 180 g 89 t 3 others

ORIGIN

Query Match

Best Local Similarity 32.7%; Score 17; DB 9; Length 546;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

22 ATCGCTCTCGTCGCTGC 38

Db

346 ATCGCTCTCGTCGCTGC 330

RESULT 40
LOCUS AV603191 556 bp mRNA linear EST 27-NOV-2001
DEFINITION AV603191 Bos taurus kidney fetus Bos taurus cDNA clone E1K1014F12 5', mRNA sequence.
ACCESSION AV603191
VERSION AV603191.1 GI:9725517
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 556)
Takesuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H. and Sugimoto,Y.
Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs

TITLE

Nucleic Acids Res. 29 (22), E108 (2001)

JOURNAL MEDLINE 21570554
PUBMED 11713328
COMMENT Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5725
Fax: 81-248-25-5725
Email: kazusugi@cocoa.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.

FEATURES

source

1..556
/organism="Bos taurus"

/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1K1014F12"
/tissue_type="kidney"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_11b="Bos taurus kidney fetus"
/note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A was deleted from a NotI site"

BASE COUNT 127 a 150 c 187 g 89 t 3 others

ORIGIN

Query Match

Best Local Similarity 32.7%; Score 17; DB 9; Length 556;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

22 ATCGCTCTCGTCGCTGC 38

Db

359 ATCGCTCTCGTCGCTGC 343

Search completed: November 13, 2003, 10:59:31
Job time : 2025 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 13, 2003, 10:59:37 ; Search time 218 Seconds
(without alignments)
222.889 Million cell updates/sec

Title: us-10-081-935-2
Perfect score: 82
Sequence: 1 MGFHALIALVAGLASA 18

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Xgapop 6.0 , Ygapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=Frame+ p2n.model -DEV=x1h
-O=/cgn2.1/USPRO.spool/US10081935/funat.12112003.144406.3612/app.query.fasta_1.199
-DB=N.GeneSeq.19Jun03 -OMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=PCT -THR.MAX=100 -THR.MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=Pico -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10081935 @CGN.1.1.0 @runat.12112003.144406.3612 -NCPU=6 -ICPU=3
-NO.MMAP -JARGOUTDIR=-NEG_SCORES=0 -MAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N.GeneSeq.19Jun03:*
1: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneSeq/geneSeqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	61.0	450	24	ABN20845
2	47	57.3	5888	23	AAS59592
3	47	57.3	4403765	22	AA199683
4	47	57.3	4411529	22	AA199682
5	46	56.1	7788	23	AAS84289
6	46	56.1	10323	24	AB199479
7	44	53.7	1751	23	AB12109
8	44	53.7	1844	23	AB117907
9	44	53.7	1947	23	AB117909
10	44	53.7	4751	23	AB117906
11	44	53.7	4754	23	AB117908
12	44	53.7	4790	23	AB112108
13	44	53.7	10091	19	AAV69642
14	43	52.4	549	21	AAC94394
15	43	52.4	549	21	AAC95143
16	43	52.4	1554	25	ABZ66799
17	43	52.4	1758	23	AAS90583
18	43	52.4	6783	23	ABL03546
19	43	52.4	10035	25	ABZ66813
20	43	52.4	11115	23	ABL50562
21	43	52.4	11220	23	AAS87866
22	43	52.4	56609	21	AAAB1459
23	43	52.4	349980	21	AAF21609
24	43	52.4	349980	21	AAF21610
25	43	52.4	1437668	21	AAAB1490
26	42	51.2	498	22	AAH53816
27	42	51.2	504	24	ABN21496
28	42	51.2	806	23	AAS85360
29	42	51.2	881	24	ABO79615
30	42	51.2	951	23	AAS83866
31	42	51.2	1031	24	ABO79606
32	42	51.2	1201	24	ABO79607
33	42	51.2	1281	24	ABO79605
34	42	51.2	1474	24	ABO79602
35	42	51.2	1569	23	ABL02695
36	42	51.2	1714	24	ABO79601
37	42	51.2	1776	15	AAO58017
38	42	51.2	1824	24	ABO79594
39	42	51.2	1874	24	ABO79597
40	42	51.2	2003	24	ABO79596
41	42	51.2	2027	24	ABO79595
42	42	51.2	2053	24	ABO79598
43	42	51.2	2211	15	AAO58012
44	42	51.2	2211	15	AAO58012
45	42	51.2	2211	21	AAAS5409

ALIGNMENTS

RESULT 1
ABN20845
ABN20845 standard; CDNA; 450 BP.
ID
XX
AC
XX
XX
DT
XX
XX
DE
XX
XX
Human ORFX polynucleotide sequence SEQ ID NO:10167.
24-JUN-2002 (first entry)
Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
degenerative disorder; osteoarthritis; neurodegenerative disorder;
cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
hypertension; hypothyroidism; cholesterol ester storage disease;
immune deficiency; immune disorder; infectious disease;
autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;


```

RESULT 3
AA199683
ID AA199683 standard; DNA: 4403765 BP.
XX
AC AA199683;
XX
DT 15-JAN-2002 (first entry)
XX
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
XX variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS Mycobacterium tuberculosis.
XX
US6294328-B1.
XX
PD 25-SEP-2001.
XX
PF 24-JUN-1998; 98US-0103840.
XX
PR 24-JUN-1998; 98US-0103840.
XX
PA (GENO-) INST GENOMIC RES.
XX
PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX
DR WPI, 2001-647261/74.
XX
PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ -
XX
PS Claim 4; SEQ ID NO 2; 3pp + Sequence listing; English.
XX
XX The invention relates to evaluating strain variation within and between
XX different populations of the tuberculosis bacterial pathogen,
XX Mycobacterium tuberculosis or related Mycobacterium by determining the
XX nucleotide sequence of the first strain at positions in the complete
XX sequence of the genome that correspond to positions that differ in the
XX nucleotide sequences of M. tuberculosis strains CDC 1551 (AA199683) and
XX H37Rv (AA199682). The method is useful for evaluating strain variation of
XX M. tuberculosis and has valuable application in the fields of
XX tuberculosis genetics, epidemiology, patient treatment and epidemic
XX monitoring.
XX CC Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other
XX
Alignment Scores:
Pred. No.: 8,89e+05 Length: 4403765
Score: 47.00 Matches: 9
Percent Similarity: 92.86% Conservative: 4
Best Local Similarity: 64.29% Mismatch: 1
Query Match: 57.32% Indels: 0
DB: 22 Gaps: 0
US-10-081-935-2 (1-18) x AA199683 (1-4403765)
OY 3 PheserHiaLeuIleALeuValaAaIaGilyLeuA 16
Db 1218600 TTCCGCCACCGCCTGTGTCGATCGTTGCCCTGAATTGCGC 1218641
RESULT 4
AA199682
ID AA199682 standard; DNA: 4411529 BP.
XX
AC AA199682;
XX
DT 15-JAN-2002 (first entry)

```

```

DE   Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX   Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KM   variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX   Mycobacterium tuberculosis.
OS   Mycobacterium tuberculosis.
XX   US6294328-B1.
PN   25-SEP-2001.
PD   24-JUN-1998; 98US-0103840.
PF   24-JUN-1998; 98US-0103840.
PR   24-JUN-1998; 98US-0103840.
PA   (GENO-) INST GENOMIC RES.
PI   Fleischmann RD, White OR, Fraser CM, Venter JC;
XX   WPI; 2001-647261/74.
XX   Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT   determining the nucleotide sequence of the strain at positions in the
PT   genome corresponding to positions where M. tuberculosis strains CDC
PT   1551 and H37Rv differ -
XX   Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX   The invention relates to evaluating strain variation within and between
XX   different populations of the tuberculosis bacterial pathogen,
XX   Mycobacterium tuberculosis or related Mycobacterium by determining the
XX   nucleotide sequence of the first strain at positions in the complete
XX   sequence of the genome that correspond to positions that differ in the
XX   nucleotide sequences of M. tuberculosis strains CDC 1551 (AA199683) and
XX   H37Rv (AA199682). The method is useful for evaluating strain variation of
XX   M. tuberculosis and has valuable application in the fields of
XX   tuberculosis genetics, epidemiology, patient treatment and epidemic
XX   monitoring.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?docid=6294328B1.
XX   SO   Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;

Alignment Scores:
Pred. No.:      8.9e+05      Length:      4411529
Score:          47.00        Matches:      9
Percent Similarity: 92.86%    Conservative: 4
Best Local Similarity: 64.29%  Mismatches:  1
Query Match:     57.32%      Indels:      0
DB:              22          Gaps:        0

US-10-081-935-2 (1-18) x AA199682 (1-4411529)

Oy   3 PheSerHisAlaLeuIleAlaLeuValIleAlaGlyLeuAla 16
      |||:::|||||:::|||||:::|||||:::|||||
Db   1219072 TTCGCGACAGCCCTGTGTCATCTGCTCCCTGAATTGGC 1219113

RESULT 5
AAS844289
ID   AAS84289 standard; cDNA; 7788 BP.
XX
XX   AAS84289;
AC
XX
XX   13-FEB-2002 (first entry)
DE   DNA encoding novel human diagnostic protein #20093.
XX
XX   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX   Homo sapiens.
OS

```

Alignment Scores: 1004 C; 980 G; 1457 T; 0 other;

Pred. No.: 833 Length: 4754
Score: 44.00 Matches: 10
Percent Similarity: 80.00% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 3
Query Match: 53.66% Indels: 0
DB: 23 Gaps: 0

US-10-081-935-2 (1-18) x ABL12108 (1-4754)

OY 4 SerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
DB 2107 GCACATGCCCGCATTCCTTCTACTTGCAGCCAGCCAGCC 2151

RESULT 12

ABL12108
ID ABL12108 standard; cDNA; 4790 BP.

XX ABL12108;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 30806.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Vencer JC, Adams M, Li PWD, Myers EW;

XX WPI, 2001-656860/75.

XX P-PSDB; ABB68005.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Claim 1; SEQ ID NO 30806; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL16175) and the encoded proteins

XX (ABB5737-ABB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pcc_sequences.

XX Sequence 4790 BP; 1324 A; 1012 C; 986 G; 1468 T; 0 other;

XX SQ

Alignment Scores:

Pred. No.: 841 Length: 4790
Score: 44.00 Matches: 10
Percent Similarity: 80.00% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 3
Query Match: 53.66% Indels: 0
DB: 23 Gaps: 0

US-10-081-935-2 (1-18) x ABL12108 (1-4790)

OY 4 SerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
DB 2107 GCACATGCCCGCATTCCTTCTACTTGCAGCCAGCCAGCC 2151

RESULT 13

AAV69642
ID AAV69642 standard; DNA; 10091 BP.

XX AAV69642;

XX 02-FEB-1999 (first entry)

XX Drosophila fat facets related Y (DFFRY) gene sequence.

XX Non-recombining region; human; Y chromosome; X homologue; testis;

XX infertility; sperm; Drosophila fat facets related Y; inhibitor;

XX gene alteration; DFRY, ss.

XX Homo sapiens.

XX WO9846747-A2.

XX 22-OCT-1998.

XX 10-APR-1998; 98WO-US07115.

XX 11-APR-1997; 97US-0041877.

XX (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX Lahn BT, Page DC;

XX WPI, 1998-568729/48.

XX P-PSDB; AAW61512.

XX Novel genes in the non-combining region of Y chromosome - useful to

XX diagnose if male infertility or reduced sperm count has a genetic

XX basis

XX Claim 7; Fig 7A-D; 54pp; English.

XX This represents the nucleotide sequence of Drosophila fat facets related

XX CC Y (DFFRY) gene. The invention relates to genes occurring on the non-

XX CC recombining region of the human Y chromosome. The sequences fall

XX CC into two classes: (1) X-homologous DNA which are expressed in many

XX CC organs, having functional X homologues and (2) testis-specific DNA

XX CC sequences. Y chromosomal DNA from males with known conditions such as

XX CC infertility and reduced sperm count can be assessed using the invention

XX CC to determine whether the condition is associated with or caused by the

XX CC occurrence of the gene or gene alteration. Candidate inhibitors of the

XX CC enzymatic activity of the genes can be assessed using in vitro assays.

XX Sequence 10091 BP; 3208 A; 1797 C; 2103 G; 2983 T; 0 other;

XX SQ

Alignment Scores:

Pred. No.: 2.09e+03 Length: 10091
Score: 44.00 Matches: 9
Percent Similarity: 76.47% Conservative: 4
Best Local Similarity: 52.94% Mismatches: 4
Query Match: 53.66% Indels: 0
DB: 19 Gaps: 0

US-10-081-935-2 (1-18) x AAV69642 (1-10091)

OY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSer 17

DB 3483 CTTCAACAAATCATGCTTGTACTTGTGTGACGAGAAACCTTGCAAC 3533

RESULT 14

AAC94394
ID AAC94394 standard; cDNA; 549 BP.

XX

XX AAC94394;
XX
DT 19-FEB-2001 (first entry)
XX
DE Cat flea head and nerve cord (HNC) cDNA, SEQ ID NO:889.
XX
XX
XX Cat flea: head and nerve cord nucleic acid; HNC; flea infestation;
XX vaccine; antiparasitic; therapeutic target; diagnosis; detection; ss.
XX
OS Ctenocephalides felis.
XX
XX WO20061621-A2.
XX
XX 19-OCT-2000.
XX
PD 07-APR-2000; 2006WO-US09437.
XX
PF 09-APR-1999; 99US-0128704.
XX
PR (HESK-) HESKA CORP.
XX
XX Brandt KS, Gaines RJ, Stinchcomb DT, Wisniewski N;
XX
XX WPI: 2000-656323/63.
XX
XX
XX The invention relates to novel cat flea (*Ctenocephalides felis*) nucleic
XX acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
XX or head and nerve cord (HNC) tissue. The invention also relates to the
XX encoded proteins. The invention additionally encompasses expression
XX constructs, recombinant viruses and recombinant cells comprising the
XX nucleic acids of the invention, recombinant production of the proteins,
XX antibodies against the proteins, a method of identifying inhibitors of
XX the proteins, and compositions comprising the inhibitors for
XX administration to an animal. The nucleic acids, and the proteins they
XX encode may be used in the prevention, treatment and diagnosis of diseases
XX associated with flea infestations. For example, the nucleic acids may be
XX used to produce an HMT or HNC protein according to standard recombinant
XX DNA methodology by inserting the nucleic acids into a host cell and
XX culturing the cell to express the protein. The HMT and HNC nucleic acids
XX may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
XX and quantitate the presence of cat flea or other homologous nucleic acid
XX sequences in samples. They may also be used to study the expression and
XX function of the proteins and their role in metabolism. The HMT and HNC
XX nucleic acids may be used as antigens in the production of specific
XX antibodies, and in assays to identify modulators (agonists and
XX antagonists) of HMT and/or HNC protein expression and activity. The
XX anti-HMT/HNC protein antibodies and antagonists may also be used to
XX downregulate protein expression and activity. The antibodies may also be
XX used as diagnostic agents for detecting the presence of flea polypeptides
XX in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
XX present sequence represents a cat flea HNC cDNA of the invention.
XX
XX
SO Sequence 549 BP, 158 A, 126 C, 137 G, 128 T, 0 other;

Alignment Scores:

Pred. NO.:	91.4	Length:	549
Score:	43.00	Matches:	9
Percent Similarity:	72.22%	Conservative:	4
Best Local Similarity:	50.00%	Mismatches:	5
Query Match:	52.44%	Indels:	0
DB:	21	Gaps:	0

US-10-081-935-2 (1-18) x AAC94394 (1-549)

0Y 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
DB 21 ATGAATAATTCGCGATGACCAATTTCGGCCCTGCTGTGGTTTGGCACTACT 74

AC	XX	AAC95143;
AC	XX	AAC95143; standard; cDNA; 549 BP.
AD	XX	
AE	XX	
AF	XX	
AG	XX	
AH	XX	
AI	XX	
AJ	XX	
AK	XX	
AL	XX	
AM	XX	
AN	XX	
AO	XX	
AP	XX	
AQ	XX	
AR	XX	
AS	XX	
AT	XX	
AV	XX	
AW	XX	
AX	XX	
AY	XX	
AZ	XX	
BA	XX	
BB	XX	
BC	XX	
BD	XX	
BE	XX	
BF	XX	
BG	XX	
BH	XX	
BI	XX	
BJ	XX	
BK	XX	
BL	XX	
BM	XX	
BN	XX	
BO	XX	
BP	XX	
BQ	XX	
BR	XX	
BS	XX	
BT	XX	
BV	XX	
BW	XX	
BX	XX	
BY	XX	
BZ	XX	
CA	XX	
CB	XX	
CC	XX	
CD	XX	
CE	XX	
CF	XX	
CG	XX	
CH	XX	
CI	XX	
CJ	XX	
CK	XX	
CL	XX	
CM	XX	
CN	XX	
CO	XX	
CP	XX	
CQ	XX	
CR	XX	
CS	XX	
CT	XX	
CU	XX	
CV	XX	
CW	XX	
CX	XX	
CY	XX	
CZ	XX	
DA	XX	
DB	XX	
DC	XX	
DD	XX	
DE	XX	
DF	XX	
DG	XX	
DH	XX	
DI	XX	
DJ	XX	
DK	XX	
DL	XX	
DM	XX	
DN	XX	
DO	XX	
DP	XX	
DQ	XX	
DR	XX	
DS	XX	
DT	XX	
DU	XX	
DV	XX	
DW	XX	
DX	XX	
DY	XX	
DZ	XX	
EA	XX	
EB	XX	
EC	XX	
ED	XX	
EE	XX	
EF	XX	
EG	XX	
EH	XX	
EI	XX	
EJ	XX	
EK	XX	
EL	XX	
EM	XX	
EN	XX	
EO	XX	
EP	XX	
EQ	XX	
ER	XX	
ES	XX	
ET	XX	
EU	XX	
EV	XX	
EW	XX	
EX	XX	
EY	XX	
EZ	XX	
FA	XX	
FB	XX	
FC	XX	
FD	XX	
FE	XX	
FF	XX	
FG	XX	
FH	XX	
FI	XX	
FJ	XX	
FK	XX	
FL	XX	
FM	XX	
FN	XX	
FO	XX	
FP	XX	
FQ	XX	
FR	XX	
FS	XX	
FT	XX	
FU	XX	
FV	XX	
FW	XX	
FX	XX	
FY	XX	
FZ	XX	
GA	XX	
GB	XX	
GC	XX	
GD	XX	
GE	XX	
GF	XX	
GG	XX	
GH	XX	
GI	XX	
GJ	XX	
GK	XX	
GL	XX	
GM	XX	
GN	XX	
GO	XX	
GP	XX	
GQ	XX	
GR	XX	
GS	XX	
GT	XX	
GU	XX	
GV	XX	
GW	XX	
GX	XX	
GY	XX	
GZ	XX	
HA	XX	
HB	XX	
HC	XX	
HD	XX	
HE	XX	
HF	XX	
HG	XX	
HH	XX	
HI	XX	
HJ	XX	
HK	XX	
HL	XX	
HM	XX	
HN	XX	
HO	XX	
HP	XX	
HQ	XX	
HR	XX	
HS	XX	
HT	XX	
HU	XX	
HV	XX	
HW	XX	
HX	XX	
HY		

US-10-081-935-2 (1-18) x AAC95143 (1-549)
 QY 1 MetGlnPheSerIslaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
 DB 21 ATGAAATTCGACGATAGCAATTTGGCGCTGCTGTGCTTGGCATCAGCT 74

RESULT 16

ID ABZ66799 standard; DNA; 1554 BP.

AC ABZ66799;

DT 21-MAR-2003 (first entry)

DE Orthosomycin biosynthetic polynucleotide SEQ ID NO 260.

KW Orthosomycin; biosynthesis; evernimicin; avilamycin; gene; ds.

OS Micromonospora carbonacea africana.

PN WO200279505-A2.

PD 10-OCT-2002.

PF 28-MAR-2002; 2002WO-CA00432.

PR 28-MAR-2001; 2001US-279095P.

PR 30-MAR-2001; 2001US-279709P.

PR 20-APR-2001; 2001US-285214P.

PA (ECOP-) ECOP1A BIOSCIENCES INC.

PI Farnet CM, Zazopoulos E, Staffa A;

DR WPI: 2003-058435/05.

DR P-PSDB; ABP99336.

PT Identifying orthosomycin biosynthetic gene, gene fragment or gene

PT cluster, by detecting presence of nucleic acid sequence corresponding

PT to 17 of flamhamycins protein families -

XX Example 1; Page 409; 511pp; English.

XX The invention relates to identifying orthosomycin biosynthetic genes and

XX its fragment/gene cluster (ABZ66799-ABZ666813), comprising detecting

XX the presence of a nucleic acid sequence coding for a polypeptide

XX (ABP99207-ABP99362). The method is useful for identifying an orthosomycin

XX biosynthetic gene, gene fragment or gene cluster, especially an

XX evenimicin-type or avilamycin-type orthosomycin biosynthetic gene,

XX gene fragment or gene cluster. The method is useful for detecting the

XX presence of any organism that contains DNA for the production of

XX orthosomycins (both evenimicin-type orthosomycins and avilamycin-type

XX orthosomycins) regardless of the level at which genes for orthosomycin

XX production are expressed by the organism or the amount of orthosomycin

XX produced by the organism. This allows for the detection of new

XX orthosomycin natural products, not produced by the organism.

SQ Sequence 1554 BP; 267 A; 615 C; 474 G; 198 T; 0 other;

Alignment Scores:

Pred. No.: 326 Length: 1554

Score: 43.00 Matches: 12

Percent Similarity: 63.64% Conservative: 2

Best Local Similarity: 54.55% Mismatches: 2

Query Match: 52.44% Indels: 6

DB: 25 Gaps: 1

US-10-081-935-2 (1-18) x ABZ66799 (1-1554)

QY 1 MetGlnPheSerIslaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 14
 DB 1 ATGAAATTCGACGATAGCAATTTGGCGCTGCTGTGCTTGGCATCAGCT 60

QY 15 LeuAla 16
 DB 61 CTCGCC 66

RESULT 17

ID AAS90583 standard; cDNA; 1758 BP.

AC AAS90583;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #26387.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PSDB; ABG26396.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX Claim 1; SEQ ID NO 26387; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful in medical

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits to assess biodiversity

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. AAS64197-AAS94564 represent novel human

XX diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1758 BP; 653 A; 368 C; 391 G; 346 T; 0 other;

Alignment Scores:

Pred. No.: 379 Length: 1758

Score: 43.00 Matches: 8

Percent Similarity: 80.00% Conservative: 4

Best Local Similarity: 53.33% Mismatches: 3

Query Match: 52.44% Indels: 0

DB: 23 Gaps: 0

AC AAA81459;
 XX 04-DEC-2000 (first entry)
 XX N. meningitidis partial DNA sequence gnm_7 SEQ ID NO:7.
 DE
 XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 XX antigen; vaccine; diagnosis; infection; antibacterial; identification;
 XX Meningococcus B; MenB; ds.
 XX Neisseria meningitidis.
 OS
 XX WO20002430-A2.
 PN
 XX 20-APR-2000.
 PD
 XX 08-OCT-1999; 99WO-US23573.
 PF
 XX 09-OCT-1999; 98US-0103794.
 PR 30-APR-1999; 99US-0132068.
 XX
 XX (CHIR) CHIRON CORP.
 PA
 XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC,
 PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Pizza M;
 XX
 XX WPI: 2000-318079/27.
 DR
 XX Isolated nucleotide sequences of Neisseria meningitidis which can be
 PT used in the diagnosis and treatment of N. meningitidis infection and
 PT other Neisserial infections, for example, N.gonorrhoea -
 XX
 XX
 PS Claim 7: Page 303-320; 1760pp; English.
 XX
 XX The present invention describes methods of obtaining immunogenic
 CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
 CC represent specifically claimed Neisseria meningitidis genomic DNA
 CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
 CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
 CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
 CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
 CC sequences, which are all used in the exemplification of the present
 CC invention. The nucleic acid sequences, protein sequences, and antibodies
 CC against them, can be used in the manufacture of a composition. The
 CC composition can be used as a medicament (or in the manufacture of a
 CC medicament) for treating, preventing or diagnosing infection due to
 CC Neisserial bacteria. For example, some of the identified proteins could
 CC be components of vaccines against Meningococcus B; against all serotypes;
 CC and/or against all pathogenic Neisseriae. Identification of sequences
 CC from the bacterium will also facilitate production of biological probes,
 CC particularly organism-specific probes. Attempts to make efficacious
 CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
 CC Multivalent vaccines have also been tried but none have successfully
 CC overcome antigenic variability. The provision of further, complete
 CC sequences may provide an opportunity to identify secreted or surface
 CC exposed proteins that may be presumed targets for the immune system and
 CC which are not antigenically variable or at least more conserved than
 CC other more variable regions.
 XX
 SO Sequence 56609 BP; 14040 A; 15839 C; 14415 G; 12313 T; 2 other;

Alignment Scores:
 Pred. No.: 2.65e+04 Length: 56609
 Score: 43.00 Matches: 9
 Percent Similarity: 92.31% Conservative: 3
 Best Local Similarity: 69.23% Mismatches: 1
 Query Match: 52.44% Indels: 0
 DB: 21 Gaps: 0

US-10-081-935-2 (1-18) x AAA81459 (1-56609)

QY 6 AAlaenllAalauValAlaAaglyLeuAlaSerAla 18
 |||||:::|||||:::|||||:::|||||:::|||||
 Db 8690 GCCTTCATCTGATGCTCGCGCTCGCGCT 8652
 RESULT 23
 AAF21609
 ID AAF21609 standard; DNA; 349980 BP.
 XX
 XX AAF21609;
 AC
 XX 13-MAR-2001 (first entry)
 DT
 XX Neisseria meningitidis B nucleotide sequence SEQ ID NO:110.
 DE
 XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
 XX diagnosis; antigen; detection; infection; gene therapy; antibacterial;
 XX ds.
 XX Neisseria meningitidis.
 OS
 XX WO200066791-A1.
 PN
 XX 09-NOV-2000.
 PD
 XX 08-MAR-2000; 2000MO-US05928.
 PF
 XX 30-APR-1999; 99US-0132068.
 PR 08-OCT-1999; 99WO-US23573.
 PR 28-FEB-2000; 2000GB-0004695.
 XX
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
 PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
 PI Frazer CM, Grandi G;
 XX
 XX WPI: 2000-647603/62.
 DR
 XX Neisseria meningitidis B full length genome sequence and open reading
 PT frames are used to detect, treat and prevent Neisserial infections -
 PT
 XX
 XX Claim 7; Appendix A; 692pp; English.
 XX
 XX The present invention describes the full length genome of
 CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
 CC to AAF21613 represent fragments of the NMB genomic sequence, as the
 CC sequence was too long to go in a record on its own it was split into 8
 CC sequence which overlap each other at the beginning and end of each
 CC the beginning of AAF21607, the last 49980 bp of AAF21544 is repeated at
 CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
 CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
 CC AAF21606 represent PCR primers which are used in the exemplification of
 CC the present invention. The NMB genome and fragments from it have
 CC antibacterial activity, and can be used in vaccines and gene therapy.
 CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
 CC proteins can be used in compositions for treating or preventing infection
 CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
 CC presence of Neisserial bacteria or of antibodies raised to Neisserial
 CC bacteria. Computers, computer memory, computer storage medium or computer
 CC databases can be used in a search to identify open reading frames (ORFs)
 CC or coding sequences within the NMB genome. The DNA sequences provide
 CC further opportunities to find antigenic or immunogenic proteins which are
 CC more effective in vaccines than the outer membrane proteins currently
 CC used.
 XX
 SO Sequence 349980 BP; 81351 A; 86755 C; 95584 G; 86290 T; 0 other;

Alignment Scores:
 Pred. No.: 2.4e+05 Length: 349980
 Score: 43.00 Matches: 9
 Percent Similarity: 92.31% Conservative: 3

Best Local Similarity: 69.23% Mismatches: 1
Query Match: 52.44% Indels: 0
DB: 21 Gaps: 0

US-10-081-935-2 (1-18) x AAF21609 (1-349980)

OY 6 A1aleu1ealaleuVal1Ala1aG1yleu1a1aSer1a1a 18
DB 317275 GCCTGCTCATTGATGTCGCGGCTCGCTCGCGCT 317313

RESULT 24

AAF21610
ID AAF21610 standard; DNA; 349980 BP.

XX AAF21610;

DT 13-MAR-2001 (first entry)

DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:111.

XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KM diagnosis; antigen; detection; infection; gene therapy; antibacterial;
ds.

XX Neisseria meningitidis.

XX WO20006791-A1.

XX 09-NOV-2000.

PF 08-MAR-2000; 2000WO-US05928.

XX 30-APR-1999; 99US-0132068.

PR 08-OCT-1999; 99WO-US23573.

PR 28-FEB-2000; 2000GB-0004695.

XX (CHIR) CHIRON CORP.

PA (GENO-) INST GENOMIC RES.

PI Pizze M, Hickey E, Peterson J, Tettein H, Venter JC, Masignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;

XX WPI; 2000-647603/62.

XX Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
XX

PS Claim 7; Appendix A; 692pp; English.

XX The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC the beginning of AAF21607, the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or as antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.

SO Sequence 349980 BP; 86771 A; 92803 C; 86340 G; 84066 T; 0 other;
Alignment Scores: 2.4e+05 Length: 349980
Pred. No.: 43.00 Matches: 9
Score: 92.31% Percent Similarity: 3
Best Local Similarity: 69.23% Mismatches: 1
Query Match: 52.44% Indels: 0
DB: 21 Gaps: 0

US-10-081-935-2 (1-18) x AAF21610 (1-349980)

OY 6 A1aleu1ealaleuVal1Ala1aG1yleu1a1aSer1a1a 18
DB 17275 GCCTGCTCATTGATGTCGCGGCTCGCTCGCGCT 17313

RESULT 25

AAAB1490
ID AAAB1490 standard; DNA; 1437668 BP.

XX AAAB1490;

DT 04-DEC-2000 (first entry)

DE N. meningitidis B full length genome DNA sequence SEQ ID NO:1068.

XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KM antigen; vaccine; diagnosis; infection; antibacterial; identification;
KM Meningococcus B; Memb; ds.

XX Neisseria meningitidis.

XX WO200022430-A2.

XX 20-APR-2000.

PF 08-OCT-1999; 99WO-US23573.

XX 09-OCT-1998; 98US-0103794.

PR 30-APR-1999; 99US-0132068.

XX (CHIR) CHIRON CORP.

PI Frazer CM, Hickey E, Peterson J, Tettein H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizze M;

XX WPI; 2000-318079/27.

XX Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX

PS Claim 7; Page 866-1272; 1760pp; English.

XX The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAAB1453 to AAAB2414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAAB1260 to AAAB1303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAAB1254 to
CC AAAB1259 and AAAB1304 to AAAB1321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAAB1322 to
CC AAAB1452 represent Neisseria meningitidis Memb polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neisseriae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious


```

XX AC AB079607;
XX XX 25-NOV-2002 (first entry)
XX DT
XX DE T. reesei cbh1 gene fragment.
XX XX Secreted protein; transcription; secretion stress; fungal; cbh1;
XX KM promoter; ds.
XX OS Trichoderma reesei.
XX PN WO200264624-A2.
XX PD 22-AUG-2002.
XX PF 13-FEB-2002; 2002WO-FI00116.
XX PR 13-FEB-2001; 2001FI-0000272.
XX PA (VALM ) VALTION TEKNIILINEN TUTKIMUSKESKUS.
XX PI Pakula T, Saloheimo M, Uusitalo J, Huuskonen A, Watson A;
XX PI Jeenes D, Archer D, Penttilae M;
XX DR WPI; 2002-657583/70.
XX PT New DNA sequence located in a promoter of a secreted protein, useful
XX PT for optimized protein production and design better fungal strains for
XX PT protein production
XX PS Disclosure; Page 81; 84pp; English.
XX CC The invention relates to a DNA sequence located in a promoter of a
XX CC secreted protein, where the DNA sequence mediates transcriptional down-
XX CC regulation of secreted proteins under secretion stress. The DNA sequence,
XX CC promoter or the fungal host is useful to optimize protein production. The
XX CC methods are useful to or to modify fungal strains for other purposes and
XX CC selectively regulate the expression of certain undesired or desired
XX CC proteins in the host. The present sequence represents a T. reesei DNA
XX CC fragment located upstream of the cbh1 gene promoter.
XX SQ Sequence 1201 BP; 328 A; 273 C; 301 G; 299 T; 0 other;

Alignment Scores:
Pred. No.: 366 Length: 1201
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 51.22% Indels: 0
DB: 24 Gaps: 0

US-10-081-935-2 (1-18) x AB079607 (1-1201)
QY 3 PheserHisAlaLeuIleAlaLeuValAlaIagIleu 15
Db 452 TTTAGCCACTTGATTGTACAAATTATTACCGCTGTATA 414

RESULT 33
AB079605/c
ID AB079605 standard; DNA; 1281 BP.
XX AC AB079605;
XX DT 25-NOV-2002 (first entry)
XX DE T. reesei cbh1 gene fragment.
XX KM Secreted protein; transcription; secretion stress; fungal; cbh1;
XX KM promoter; ds.
XX OS Trichoderma reesei.

```

```

PN WO200264624-A2.
XX XX 22-AUG-2002.
XX PD
XX PF 13-FEB-2002; 2002WO-FI00116.
XX PR 13-FEB-2001; 2001FI-0000272.
XX PA (VALM ) VALTION TEKNIILINEN TUTKIMUSKESKUS.
XX PI Pakula T, Saloheimo M, Uusitalo J, Huuskonen A, Watson A;
XX PI Jeenes D, Archer D, Penttilae M;
XX DR WPI; 2002-657583/70.
XX PT New DNA sequence located in a promoter of a secreted protein, useful
XX PT for optimized protein production and design better fungal strains for
XX PT protein production
XX PS Disclosure; Page 80; 84pp; English.
XX CC The invention relates to a DNA sequence located in a promoter of a
XX CC secreted protein, where the DNA sequence mediates transcriptional down-
XX CC regulation of secreted proteins under secretion stress. The DNA sequence,
XX CC promoter or the fungal host is useful to optimize protein production. The
XX CC methods are useful to or to modify fungal strains for other purposes and
XX CC selectively regulate the expression of certain undesired or desired
XX CC proteins in the host. The present sequence represents a T. reesei DNA
XX CC fragment located upstream of the cbh1 gene promoter.
XX SQ Sequence 1281 BP; 356 A; 290 C; 317 G; 318 T; 0 other;

Alignment Scores:
Pred. No.: 396 Length: 1281
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 51.22% Indels: 0
DB: 24 Gaps: 0

US-10-081-935-2 (1-18) x AB079605 (1-1281)
QY 3 PheserHisAlaLeuIleAlaLeuValAlaIagIleu 15
Db 532 TTTAGCCACTTGATTGTACAAATTATTACCGCTGTATA 494

RESULT 34
AB079602/c
ID AB079602 standard; DNA; 1474 BP.
XX AC AB079602;
XX DT 25-NOV-2002 (first entry)
XX DE T. reesei cbh1 gene fragment.
XX KM Secreted protein; transcription; secretion stress; fungal; cbh1;
XX KM promoter; ds.
XX OS Trichoderma reesei.
XX PN WO200264624-A2.
XX PD 22-AUG-2002.
XX PF 13-FEB-2002; 2002WO-FI00116.
XX PR 13-FEB-2001; 2001FI-0000272.
XX PA (VALM ) VALTION TEKNIILINEN TUTKIMUSKESKUS.
XX PI Pakula T, Saloheimo M, Uusitalo J, Huuskonen A, Watson A;
XX PI Jeenes D, Archer D, Penttilae M;

```

XX DR MPI: 2002-657583/70.
XX PT New DNA sequence located in a promoter of a secretable protein, useful
PT for optimized protein production and design better fungal strains for
PT protein production -
XX PS Disclosure; Page 79; 84pp; English.
XX CC The invention relates to a DNA sequence located in a promoter of a
CC secretable protein, where the DNA sequence mediates transcriptional down-
CC regulation of secreted proteins under secretion stress. The DNA sequence,
CC promoter or the fungal host is useful to optimize protein production. The
CC methods are useful to or to modify fungal strains for other purposes and
CC selectively regulate the expression of certain undesired or desired
CC proteins in the host. The present sequence represents a T. reesei DNA
CC fragment located upstream of the cbh1 gene promoter.
SQ Sequence 1474 BP; 397 A; 337 C; 357 G; 383 T; 0 other;

Alignment Scores:
Pred. No.: 470 Length: 1474
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 51.22% Indels: 0
DB: 24 Gaps: 0

US-10-081-935-2 (1-18) x ABQ79602 (1-1474)
QY 3 PhseRHsAlaleuilealaleuValAlaAlaGlyLeu 15
Db 1466 TTTAGCCACTTGATGTGACATTATTATAGCCGCTGTATA 1428
RESULT 35
ABL02695
ID ABL02695 standard; cDNA; 1569 BP.
XX AC ABL02695;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 2567.
XX KM Drosophila: developmental biology; cell signalling; insecticide;
XX KM pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR MPI: 2001-656860/75.
XX DR P-PSDB; ABB58592.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX PS Claim 1; SEQ ID NO 2567; 21pp + Sequence Listing; English.
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57373-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 1569 BP; 399 A; 465 C; 381 G; 324 T; 0 other;

Alignment Scores:
Pred. No.: 507 Length: 1569
Score: 42.00 Matches: 8
Percent Similarity: 91.67% Conservative: 3
Best Local Similarity: 66.67% Mismatches: 1
Query Match: 51.22% Indels: 0
DB: 23 Gaps: 0

US-10-081-935-2 (1-18) x ABL02695 (1-1569)
QY 4 SerHsAlaleuilealaleuValAlaAlaGlyLeu 15
Db 807 TCCCACTCCGCTGTTCCCTCGTCCGACACTT 842
RESULT 36
ABQ79601/c
ID ABQ79601 standard; DNA; 1714 BP.
XX AC ABQ79601;
XX DT 25-NOV-2002 (first entry)
XX DE T. reesei cbh1 gene fragment.
XX KM Secreted protein; transcription; secretion stress; fungal; cbh1;
XX KM promoter; ds.
XX OS Trichoderma reesei.
XX PN WO200264624-A2.
XX PD 22-AUG-2002.
XX PF 13-FEB-2002; 2002WO-FI00116.
XX PR 13-FEB-2001; 2001FI-0000272.
XX PA (VALM) VALTION TEKNIILLINEN TUTKIMUSKESKUS.
XX PI Pakula T, Saloheimo M, Unsitalo J, Huuskonen A, Watson A;
XX PI Jeenes D, Archer D, Penttilae M;
XX DR MPI: 2002-657583/70.
XX PT New DNA sequence located in a promoter of a secretable protein, useful
PT for optimized protein production and design better fungal strains for
PT protein production -
XX PS Disclosure; Page 78; 84pp; English.
XX CC The invention relates to a DNA sequence located in a promoter of a
CC secretable protein, where the DNA sequence mediates transcriptional down-
CC regulation of secreted proteins under secretion stress. The DNA sequence,
CC promoter or the fungal host is useful to optimize protein production. The
CC methods are useful to or to modify fungal strains for other purposes and
CC selectively regulate the expression of certain undesired or desired
CC proteins in the host. The present sequence represents a T. reesei DNA
CC fragment located upstream of the cbh1 gene promoter.
SQ Sequence 1714 BP; 459 A; 389 C; 422 G; 444 T; 0 other;

```

Alignment Scores:
Pred. No.: 565 Length: 1714
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 51.22% Indels: 0
DB: 24 Gaps: 0

US-10-081-935-2 (1-18) x AB079601 (1-1714)

OY 3 PhseSerHsAlaLeuIleAlaLeuValAlaAclyleu 15
Db 1466 TTATGCCACTTGATTGACATTATTAGCCGCTGTATA 1428

RESULT 37
AA058017/c
ID AA058017 standard; DNA; 1776 BP.
XX
AC AA058017;
XX
DT 25-MAR-2003 (updated)
DT 14-SEP-1994 (first entry)
XX
DE Sequence of altered cellobiohydrolase 1 (cbh1) promoter of pMI-27.
XX
KM Promoter; cellobiohydrolase 1; cbh1; pMI-27; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_difference 704..709
/*tag= a
/*label= altered sequence
FT misc_feature 716..748
/*tag= b
/*label= polylinker
FT misc_feature 748..749
/*tag= C
/*label= deletion site
FT misc_difference 1054..1059
/*tag= d
/*label= altered sequence
FT CDS 1774..1776
/*tag= e
/*label= start codon
XX
PN MO9404673-A1.
XX
PD 03-MAR-1994.
XX
PF 19-AUG-1993; 93MO-F100330.
XX
PR 19-AUG-1992; 92US-0932485.
XX
PA (ALKO-) ALKO OY AB.
XX
PI Ilmen MH, Nakari TH, Nevalainen KM, Onnela M, Penttilae ME;
XX DR WPI; 1994-083192/10.
XX
PT Cloning promoters active in a desired environmental condition -
PT used partic. for expression of genes in Trichoderma fungal hosts
PT in glucose-contg. medium
XX
XX Example; Figure 18B; 120pp; English.
XX
CC PM1016del10(12) was used as a PCR template to yield a 280 bp fragment
CC E (primers = AA058023 and AA058024), beginning at the promoter
CC internal polylinker and ending at -720, and a 720 bp fragment F
CC (primers = AA058026 and AA058028), beginning at -720 and ending at
CC KspI at -16. Fragments D and E were purified from agarose gel
CC and digested with BstEII-XbaI and XbaI-KspI respectively, & ligated
CC to the 7.8 kb BstEII-KspI fragment of pMI-25 to produce produce

```

```

CC cbh1 promoter of pMI-27 has sequence alterations at
CC posns. -1505-1500 (genomic sequence GTGGCG, altered sequence
CC TCTAAA) and at posn. -720 to -715 (genomic sequence GTGGCG, altered
CC sequence TCTTGA) upstream of the translation initiation codon of intact
CC cbh1 promoter. The sequence of the altered cbh1 promoter in pMI-27
CC is provided in AA058017.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 1776 BP; 490 A; 399 C; 430 G; 457 T; 0 other:

Alignment Scores:
Pred. No.: 590 Length: 1776
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 51.22% Indels: 0
DB: 15 Gaps: 0

US-10-081-935-2 (1-18) x AA058017 (1-1776)

OY 3 PhseSerHsAlaLeuIleAlaLeuValAlaAclyleu 15
Db 1029 TTATGCCACTTGATTGACATTATTAGCCGCTGTATA 991

RESULT 38
AA058018/c
ID AA058018 standard; DNA; 1776 BP.
XX
AC AA058018;
XX
DT 25-MAR-2003 (updated)
DT 14-SEP-1994 (first entry)
XX
DE Sequence of altered cellobiohydrolase 1 (cbh1) promoter of pMI-28.
XX
KM Promoter; cellobiohydrolase 1; cbh1; pMI-28; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_difference 704..709
/*tag= a
/*label= altered sequence
FT misc_feature 716..748
/*tag= b
/*label= polylinker
FT misc_feature 748..749
/*tag= C
/*label= deletion site
FT misc_difference 773..778
/*tag= d
/*label= altered sequence
FT misc_difference 1054..1059
/*tag= e
/*label= altered sequence
FT CDS 1774..1776
/*tag= f
/*label= start codon
XX
PN MO9404673-A1.
XX
PD 03-MAR-1994.
XX
PF 19-AUG-1993; 93MO-F100330.
XX
PR 19-AUG-1992; 92US-0932485.
XX
PA (ALKO-) ALKO OY AB.
XX
PI Ilmen MH, Nakari TH, Nevalainen KM, Onnela M, Penttilae ME;
XX DR WPI; 1994-083192/10.
XX

```

PT Cloning promoters active in a desired environmental condition -
PT used partic. for expression of genes in Trichoderma fungal hosts
PT in glucose-contg. medium
XX
PS Example; Figure 18C; 120pp; English.
XX
CC pMOIde10(2) was used as a PCR template to yield a 280 bp fragment
CC G (primers = AA058023 and AA058025), beginning at the promoter
CC internal polylinker and ending at -720, and a 720 bp fragment H
CC (primers = AA058026 and AA058028), beginning at -720 and ending at
CC Kapi at -16. Fragments G and H were purified from agarose gel
CC and digested with BstEII-XbaI and XbaI-KspI respectively, & ligated
CC to the 7.8 kb BstEII-KspI fragment of pMI-25 to produce produce
CC pMI-28. The cbh1 promoter of pMI-28 has sequence alterations at
CC posns. -1505-1500 (genomic sequence GTGGCG, altered sequence
CC TCTAAA), -1001-996 (genomic sequence GTGGCG, altered sequence
CC TCTAAA), and at posn. -720 to -715 (genomic sequence GTGGCG, altered
CC sequence TCTGAG) upstream of the translation initiation codon of intact
CC cbh1 promoter. The sequence of the altered cbh1 promoter in pMI-28
CC is provided in AA058018.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 1776 BP; 490 A; 400 C; 429 G; 457 T; 0 other;
SO
XX
Alignment Scores:
Pred. No.: 590 Length: 1776
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 51.22% Indels: 0
DB: Gaps: 0
US-10-081-935-2 (1-18) x AA058018 (1-1776)
Qy 3 PhseRHisAlaLeuIleAlaLeuValAlaGlyIleu 15
ID 1029 TTtAGCCACTTGATTGATCAATTATTTAGCCCTGTAATA 991
XX
XX AB079594; standard; DNA; 1824 BP.
XX
AC AB079594;
XX
XX 25-NOV-2002 (first entry)
XX
XX T. reesei cbh1 gene fragment.
XX
XX DE Secreted protein; transcription; secretion stress; fungal; cbh1;
XX KW promoter; ds.
XX
XX OS Trichoderma reesei.
XX
XX OS WO200264624-A2.
XX
XX PD 22-AUG-2002.
XX
XX PF 13-FEB-2002; 2002WO-FI00116.
XX
XX PR 13-FEB-2001; 2001FI-0000272.
XX
XX (VALW) VALTION TEKNIILLINEN TUTKIMUSKESKUS.
XX
XX PI Pakula T, Saloheimo M, Uusitalo J, Huuskonen A, Watson A;
XX PI Jeenes D, Archer D, Penttilae M;
XX
XX DR WPI: 2002-657583/70.
XX
XX PT New DNA sequence located in a promoter of a secretable protein, useful
XX PT for optimized protein production and design better fungal strains for
XX PT protein production -
XX
XX PS Disclosure; Page 74; 84pp; English.

XX
CC The invention relates to a DNA sequence located in a promoter of a
CC secretable protein, where the DNA sequence mediates transcriptional down-
CC regulation of secreted proteins under secretion stress. The DNA sequence,
CC promoter or the fungal host is useful to optimize protein production. The
CC methods are useful to or to modify fungal strains for other purposes and
CC selectively regulate the expression of certain undesired or desired
CC proteins in the host. The present sequence represents a T. reesei DNA
CC fragment located upstream of the cbh1 gene promoter.
XX
SQ Sequence 1824 BP; 492 A; 406 C; 452 G; 474 T; 0 other;
SO
XX
Alignment Scores:
Pred. No.: 610 Length: 1824
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 51.22% Indels: 0
DB: Gaps: 0
US-10-081-935-2 (1-18) x AB079594 (1-1824)
Qy 3 PhseRHisAlaLeuIleAlaLeuValAlaGlyIleu 15
ID 1466 TTtAGCCACTTGATTGATCAATTATTTAGCCCTGTAATA 1428
XX
XX AB079597; standard; DNA; 1874 BP.
XX
AC AB079597;
XX
XX 25-NOV-2002 (first entry)
XX
XX DE T. reesei cbh1 gene fragment.
XX
XX KW Secreted protein; transcription; secretion stress; fungal; cbh1;
XX KW promoter; ds.
XX
XX OS Trichoderma reesei.
XX
XX OS WO200264624-A2.
XX
XX PD 22-AUG-2002.
XX
XX PF 13-FEB-2002; 2002WO-FI00116.
XX
XX PR 13-FEB-2001; 2001FI-0000272.
XX
XX (VALW) VALTION TEKNIILLINEN TUTKIMUSKESKUS.
XX
XX PI Pakula T, Saloheimo M, Uusitalo J, Huuskonen A, Watson A;
XX PI Jeenes D, Archer D, Penttilae M;
XX
XX DR WPI: 2002-657583/70.
XX
XX PT New DNA sequence located in a promoter of a secretable protein, useful
XX PT for optimized protein production and design better fungal strains for
XX PT protein production -
XX
XX PS Disclosure; Page 76; 84pp; English.
XX
XX CC The invention relates to a DNA sequence located in a promoter of a
XX CC secretable protein, where the DNA sequence mediates transcriptional down-
XX CC regulation of secreted proteins under secretion stress. The DNA sequence,
XX CC promoter or the fungal host is useful to optimize protein production. The
XX CC methods are useful to or to modify fungal strains for other purposes and
XX CC selectively regulate the expression of certain undesired or desired
XX CC proteins in the host. The present sequence represents a T. reesei DNA
XX CC fragment located upstream of the cbh1 gene promoter.
XX
XX SQ Sequence 1874 BP; 499 A; 420 C; 467 G; 488 T; 0 other;

Alignment Scores:

Pred. No.:	630	Length:	1874
Score:	42.00	Matches:	7
Percent Similarity:	84.62%	Conservative:	4
Best Local Similarity:	53.85%	Mismatches:	2
Query Match:	51.22%	Indels:	0
DB:	24	Gaps:	0

US-10-081-935-2 (1-18) x AB079597 (1-1874)

QY	3	PhSeRtHsAlaLeuIleAlaLeuValAlaIaGlyLeu	15
Db	1466	TTTAGCCACTTGATTGTAATTAATTAGCCCGCTGTATA	1428

Search completed: November 13, 2003, 12:08:30
 Job time : 842 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 13, 2003, 11:51:07 ; Search time 52 Seconds
(without alignments)
152.786 Million cell updates/sec

Title: US-10-081-935-2
Perfect score: 82
Sequence: 1 MQFSHALIALVAAGLASA 18

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xih
-Q=/cgn2_1/USPTO.spool/US10081935/runat_12112003_144407_3656/app_query.fasta_1.199
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFM=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10081935 @CGN_1_1_44 @runat_12112003_144407_3656 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	57.3	4403765	3	US-09-103-840A-2
2	47	57.3	4411529	3	US-09-103-840A-1
3	44	53.7	248	1	US-08-525-697-3
4	44	53.7	1302	1	US-08-525-697-1
5	44	53.7	10091	3	US-09-058-489-34
6	42	51.2	471	4	US-09-252-991A-8612
7	42	51.2	504	4	US-09-134-001C-959
8	42	51.2	1781	2	US-08-389-564B-21
9	42	51.2	1781	2	US-08-389-564B-22
10	42	51.2	1781	3	US-08-466-047B-21
11	42	51.2	1781	3	US-08-466-047B-22
12	42	51.2	2218	2	US-08-389-564B-18

C 13	42	51.2	2218	3	US-09-066-597-1	Sequence 1, Appli
C 14	42	51.2	2218	3	US-08-466-047B-18	Sequence 18, Appli
C 15	42	51.2	2266	2	US-08-389-564B-20	Sequence 20, Appli
C 16	42	51.2	2266	3	US-08-466-047B-20	Sequence 20, Appli
C 17	42	51.2	4176	3	US-09-254-733-1	Sequence 1, Appli
C 18	42	51.2	4176	3	US-08-374-077C-1	Sequence 1, Appli
C 19	42	51.2	8075	3	US-08-895-590-1	Sequence 1, Appli
C 20	42	51.2	8075	4	US-09-539-879A-1	Sequence 1, Appli
C 21	42	51.2	8096	3	US-09-058-489-33	Sequence 33, Appli
C 22	41	50.0	528	4	US-09-252-991A-9137	Sequence 9137, Ap
C 23	41	50.0	972	4	US-09-252-991A-15702	Sequence 15702, A
C 24	41	50.0	975	4	US-09-252-991A-15674	Sequence 15674, A
C 25	41	50.0	1056	4	US-09-252-991A-8728	Sequence 8728, Ap
C 26	41	50.0	1218	4	US-09-552-322-3	Sequence 3, Appli
C 27	41	50.0	1230	4	US-09-328-352-1918	Sequence 1918, Ap
C 28	41	50.0	1485	4	US-09-252-991A-15656	Sequence 15656, A
C 29	41	50.0	1566	4	US-09-252-991A-9220	Sequence 9220, Ap
C 30	41	50.0	1710	4	US-09-252-991A-9057	Sequence 9057, Ap
C 31	41	50.0	2123	4	US-09-724-864-18	Sequence 18, Appli
C 32	41	50.0	5045	3	US-09-390-721-1	Sequence 1, Appli
C 33	41	50.0	5045	3	US-09-390-721-3	Sequence 3, Appli
C 34	41	50.0	36412	4	US-08-311-731A-132	Sequence 132, App
C 35	40	48.8	304	4	US-08-313-294A-2891	Sequence 2891, Ap
C 36	40	48.8	465	4	US-09-252-991A-16275	Sequence 16275, A
C 37	40	48.8	504	4	US-09-512-342-19	Sequence 19, Appli
C 38	40	48.8	669	4	US-09-252-991A-3703	Sequence 3703, Ap
C 39	40	48.8	717	3	US-08-913-014A-7	Sequence 7, Appli
C 40	40	48.8	765	4	US-09-252-991A-9931	Sequence 9931, Ap
C 41	40	48.8	795	3	US-08-913-014A-8	Sequence 8, Appli
C 42	40	48.8	951	4	US-09-252-991A-16386	Sequence 16386, A
C 43	40	48.8	1053	4	US-09-252-991A-6421	Sequence 6421, Ap
C 44	40	48.8	1092	4	US-09-252-991A-16212	Sequence 16212, A
C 45	40	48.8	1179	4	US-09-252-991A-10081	Sequence 10081, A

ALIGNMENTS

RESULT 1

US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 1.95e+05
Score: 47.00
Percent Similarity: 92.86%
Best Local Similarity: 64.29%
Query Match: 57.32%
DB: 3
Length: 4403765
Matches: 9
Conservative: 4
Mismatches: 1
Indels: 0
Gaps: 0
US-10-081-935-2 (1-18) x US-09-103-840A-2 (1-4403765)

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0486/93
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-525-697-3

Alignment Scores:
Pred. No.: 3-04 Length: 248
Score: 44.00 Matches: 8
Percent Similarity: 77.78% Conservatives: 6
Best Local Similarity: 44.44% Mismatches: 4
Query Match: 53.66% Indels: 0
DB: 1 Gaps: 0

US-10-081-935-2 (1-18) x US-08-525-697-3 (1-248)

Qy 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
Db 24 ATGAAGCTTTCTCACATGCTCTCCAGCCTGCCAGCTGGGGGTGGCGACGGCT 77

RESULT 4
US-08-525-697-1
; Sequence 1, Application US/08525697
; Patent No. 5795764
; GENERAL INFORMATION:
; APPLICANT: Christgau, Stephan
; APPLICANT: Andersen, Lene N
; APPLICANT: Kauppinen, Sakari
; APPLICANT: Heldt-Hansen, Hans P
; APPLICANT: Dalboege, Henrik
; TITLE OF INVENTION: AN ENZYME EXHIBITING MANNANASE ACTIVITY
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57957640 No. 5795764disk of No. 5795764th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,697
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Harrington, James J.
; REGISTRATION NUMBER: 38,711
; REFERENCE/DOCKET NUMBER: 4004.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0486/93
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1302 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0486/93
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1302 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0486/93
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-525-697-3

Alignment Scores:
Pred. No.: 3-04 Length: 248
Score: 44.00 Matches: 8
Percent Similarity: 77.78% Conservatives: 6
Best Local Similarity: 44.44% Mismatches: 4
Query Match: 53.66% Indels: 0
DB: 1 Gaps: 0

US-10-081-935-2 (1-18) x US-08-525-697-3 (1-248)

Qy 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
Db 24 ATGAAGCTTCTCACATGCTCTCCAGCCTGCCAGCTGGGGGTGGCGACGGCT 77

RESULT 4
US-08-525-697-1
; Sequence 1, Application US/08525697
; Patent No. 5795764
; GENERAL INFORMATION:
; APPLICANT: Christgau, Stephan
; APPLICANT: Andersen, Lene N
; APPLICANT: Kauppinen, Sakari
; APPLICANT: Heldt-Hansen, Hans P
; APPLICANT: Dalboege, Henrik
; TITLE OF INVENTION: AN ENZYME EXHIBITING MANNANASE ACTIVITY
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57957640 No. 5795764disk of No. 5795764th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,697
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Harrington, James J.
; REGISTRATION NUMBER: 38,711
; REFERENCE/DOCKET NUMBER: 4004.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0486/93
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1302 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0486/93
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1302 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

[illegible]

/product="virion glycoprotein RL10"
/protein_id="AA000650.1"
/db_xref="GI:19881030"
/translation="MRVYTLGCTSTLELLAIGATAIHNDVCTSLDGQTRLLCKNRWNNN
SNPKPTATSGTHITCLANCTCKRYTEPLPLISVLGIYSAMGAGSPIATLIILLVIFV
IYSRERADDNGTVEDDFLAYRDLTRKKLDSHASKQNIVERIPYPCRLRDDEPP
LITNAEDDEEDNF"
7202..7900
/note="RL11 family"
/codon_start=1
/product="IgG Fc-binding glycoprotein RL11"
/protein_id="AA000651.1"
/db_xref="GI:19881031"
/translation="HKRDALTLKYVVSTAFCLISAKATPTPTDGLPRLRLDRTYTCE
TRNRTHASRLNTGHSVRYALPGHKPTQTMPLCSITDAHSTTRDVRKARFICTR
HSMTLVYVTEDMGVILTDYETGDAESFYFIFSRCAVEAYATCAFWKLDNRQCT
LKPGATNPDPNDLRDMWLNI SHICAWIAGLLIIIVASLVGFLRLRRVGVQNAVYRLSN
RDEPLLOFDAYSE"
7939..9021
/note="RL11 family"
/codon_start=1
/product="glycoprotein RL12"
/protein_id="AA000652.1"
/db_xref="GI:19881032"
/translation="MSETCSVATLLVLLALSOLTEEDNTNSSVNTSTNDSNISFTT
TSSPKATSTPSPFTNLTSTSYLNSVNSTHTPTLTTLPTSTTTHSFQNTIIT
EYSTYUNCITETVSSNSTESITATSCNTANTIFDNTITSTAIQTITTEIVIKAW
TTHVKESENTVLHNSVTFQNCRTVMNHPYNIITHKMRKENKPHLCEASSGHSVHK
NHKLCLQCTKNLTLFDLMVHNSGKIYAECDGQNHQAQGFILTVNATHITNTRVC
QPKDVTTPRLSTHLFNQNSYOGHYPCASANSRGMAGVGLFLMACVLLFFARRVY
KKKYRPLRDVSESEFVRYTPEHD"
9123..10124
/note="RL11 family"
/codon_start=1
/product="glycoprotein RL13"
/protein_id="AA000653.1"
/db_xref="GI:19881033"
/translation="MVDHFAITWTLVLTLAGTGVTSCHTSHSVSTTVATTSTPS
VNTNSADTSNTSTESITSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST
TTTTTTTST
HTARWTVNKTEDLCIFGPDYYSQPAICFKHQWQMTIYDVTTENAGNIVRY
HGDGHHYDKGYRLQVTSNHTTGTNRKCPNDFTSYTPDHDKNKETKTIEFPGMND
QPTFPMGHAIWAAVWVALLVMGRSSSVTVKGGKPKYKLSNKDPEYMASP"
complement(10256..10432)
/note="contains hydrophobic domain"
/codon_start=1
/product="UL2"
/protein_id="AA000654.1"
/db_xref="GI:19881034"
/translation="MRGDAVSILIVEDLDLPFGSGFNASHAYYSFRVLRGIFYVTVVV
WLVWIKLLRNFP"
10504..11097
/note="RL11 family; contains hydrophobic domain"
/codon_start=1
/product="UL4"
/protein_id="AA000655.1"
/db_xref="GI:19881035"
/translation="MRRRGVGGPGCATFKYGLLVLTSTWVSIGIHYGAYDHHVDR
RVISGNDVLTSPYHLGNSFWYKRNATGYPVCSYKTIYTFPDVVRVECKHMMNI
TIFNTEGWIYSLTESPFLDNKHYTEAFYVITDIRNNGTEBELDYDDYDYGSGF
GSGLGLESGLGLESGLSGNEETSGSGDEIQNGFLITY"
11469..11924
/note="RL11 family; contains hydrophobic domain"
/codon_start=1
/product="UL5"
/protein_id="AA000656.1"
/db_xref="GI:19881036"
/translation="MFLGYSYDFMRFGFALCRAPGHYQTSVTTLLTCAVSSSLTAS
TSAAGSTGGYRNLTORANLSSPDTPTTSLPROIAAPPAHSSGLFEOFYHGLAIHA
AWTAFVLVGLIFFANVPHVPTQTHGVHRLDRHRTLFGRRYQRLNS"
12130..12981
/note="RL11 family"
/codon_start=1

/product="glycoprotein UL6"
/protein_id="AA000657.1"
/db_xref="GI:19881037"
/translation="MPPPRSFCGLKNKWCISCLLILVAVCAPRGSPHWSRVRQT
SLSPSCSPVNCTVGGDSVLSNSTIPVACNSTEWGRYNSFWPCLQWGHMRLSGQ
NLRLSLTCSROHLTLHSTSDGYTGTYGVGNCCPHHHEIMKTCFNLVLAKPTTTAAP
TTTSVPPTLLTLTRNTVIVGSPANTVLSSTTTTAGNASVANGMEFYROYPLSLA
QTHRTAPLANATDNAEDVANLVATYASMGVLVLLLLATVLVFLDGLPQTAWRWROGR
EEQHLLL"
13052..13693
/note="RL11 family"
/codon_start=1
/product="glycoprotein UL7"
/protein_id="AA000658.1"
/db_xref="GI:19881038"
/translation="MAADPHITLMSGLCKNALWLSCLLIVAFLLASSSTTTMYVY
SGPRLPLSHANFTQDDVHMKLNETVSLCSVMVGDGVFIARSNNSLGLCPMQL
ILFMTVNDGGLYVNGTNGTTRTSTTVMFNTIGLOFAPKTTKPKKTKTKRKLSE
RTGMSLPYARDLDSVRRKDDDDNIHLGVAGLVALVIVCLMGNKLILCTH"
13799..14305
/note="RL11 family"
/codon_start=1
/product="glycoprotein UL8"
/protein_id="AA000659.1"
/db_xref="GI:19881039"
/translation="MNVTGQPPSPPPPGPSESTWNVTGGEDGERSVTTTVAVST
RPTLSGKNASTATSTAPVTTTRTTRTTRTSTSTSTSTSTSTSTSTSTSTSTSTST
LRTLSGHTTVAQPKIAVHYTWLILFVLVILFCLRIPOKLYDKWNRNRYGO
VYVDTEL"
14327..14860
/note="RL11 family"
/codon_start=1
/product="glycoprotein UL9"
/protein_id="AA000660.1"
/db_xref="GI:19881040"
/translation="WFYTFSLTYVCKNLHEVPECKVKTGYIGONVSLKSI THW
SNSTWSVHTVSSLYMCKERESNDAPFNFNVTDIQFNCTEELLLELLELTOTYAVYY
RGYDTPKPCMLYSTCYNTVKNRTYAEKHTFASSIWI PPVAVVTLIILLSVIKIPI
RIWEEWQYRYRDTVYT"
14959..15657
/note="RL11 family"
/codon_start=1
/product="glycoprotein UL10"
/protein_id="AA000661.1"
/db_xref="GI:19881041"
/translation="MPFGSYVNCGFLTWVLSAVFKYEFSEACEEHTTYETHPLLGL
GKPLLAGHOTAYWYKHKCDSTSSNDKLTPLCYYPNRSKGLKDESAENTFINYLCSN
STLIARLNLTDAGEYCKRKNSTRHDSWPTMTCYRVIVRPTKTKSKPESGPTVKR
TTYDNYSVTLAEELLFQTHRSTQNAHIAWLFIAVVIIVILFFFKIPQKLPDKWL
HKSIAKTGHFSAPH"
15758..16468
/note="RL11 family"
/codon_start=1
/product="glycoprotein UL11"
/protein_id="AA000662.1"
/db_xref="GI:19881042"
/translation="MCHRRLEFKTSFYAMICVFVSTDLQNCACVEEKARQGSNVT
FDARIYSGKNOSTYFWRGSRYNADMNFTGNLTVHSTEIRLNYTCQNFSLTLNVT
SEYKNKYTRITVSTSTNSFDTCKFLTIVIVKPKPKTKTKRQEPVTKPKTT
YVDYVYDGFATIAVSPGIVSSQRSQATTISIWICLSIIVIVVILFVFRIPQKLLP
FMCPRPKDETLMIPQTEL"
16623..18392
/codon_start=1
/product="UL13"
/protein_id="AA000663.1"
/db_xref="GI:19881043"
/translation="MVVVSRLTSYVILPLLCKVPPILLTPWDSAHGLDIDVEELSH
DYEISVVRFAVQVGMQLPAPIMGRSEIDEDEBPSSPAQADVVDVSRPPRPRRP
RLQSPYHQHQPSPSGIRWOHEELQFLQTRQELLQOREERQORQORRRYRWR
RPTFPDPDPSPYPARSSTILPARFARPSLANGAEQLNARGYGERSSQRGNASPA
SVEEYQYRVVGVNTRRRHPHRSRRNALVANGRDSLRLARLIRHQTRTFYGY
RCRYFGRNRGSRRDDGTCGAAGHGGRPGAGFSQLRERIVTDLOLFLRCHGWT
RQASRRIRTRWEEENTVMSDAASRLRAWFSRRTTYQRTWTPGENPSAEAGELAVPPA

ADGVKETSQSMVTTEGEREQTEORVEKAAKGGBETIEKIVETDTRVAETKXVE
 EQVVEEENNVVEETREERDRDDEDAIIPMGWMDLSSOLETTAAETEEVAVEK
 EETTKGADATTVTTASSETDDVITQNELPCELNNAEEISSGRAVVGTCPRREGPHR
 SFFRLCLGLWASSHARRAISVS"
 18719. .19780

CDS

Alignment Scores:

Pred. No.: 942 Length: 241087
 Score: 52.00 Matches: 10
 Percent Similarity: 93.75% Conservative: 5
 Best Local Similarity: 62.50% Mismatches: 1
 Query Match: 63.41% Indels: 0
 DB: 14 Gaps: 0

US-10-081-935-2 (1-18) x AF480884 (1-241087)

Qy 2 GlnPheSerHisAlaLeuValAlaAlaGlyLeuAlaSerAla 17

Db 66475 CGTTACGACGCGCTGGTGTGCTGCCCGCGCCACGGCAGC 66428

RESULT 6

CNS01AF9 660 bp mRNA linear PLN 02-SEP-1999
 LOCUS Botrytis cinerea strain T4 cDNA library under conditions of
 DEFINITION nitrogen deprivation.

ACCESSION AL112861

VERSION AL112861.1 GI:5827480

KEYWORDS cDNA library; nitrogen deprivation.

SOURCE Botryotinia fuckeliana

ORGANISM Botryotinia fuckeliana

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;

AUTHORS Helotiales; Sclerotiniaceae; Botryotinia.

TITLE Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.

JOURNAL Direct Submission

REFERENCE Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,

AUTHORS 78026 Versailles, France

TITLE 2 (bases 1 to 660)

JOURNAL Genoscope.

COMMENT Direct Submission

Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
 CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

The cDNA library to be analyzed within the framework of this
 project was created using a Botrytis cinerea strain which was grown
 under conditions of nitrogen deprivation, which is the normal
 situation for B. cinerea during its development on its host plant.
 The library was produced in an oriented direction, in the pBSII
 vector.

FEATURES

source Location/Qualifiers
 1. .660
 /organism="Botryotinia fuckeliana"
 /mol_type="mRNA"
 /strain="T4"
 /db_xref="taxon:40559"
 /note="Genoscope sequence ID : W05H031"

BASE COUNT 161 a 214 c 131 g 154 t

Alignment Scores:

Pred. No.: 13.6 Length: 660
 Score: 48.00 Matches: 10
 Percent Similarity: 72.22% Conservative: 3
 Best Local Similarity: 55.56% Mismatches: 5
 Query Match: 58.54% Indels: 0
 DB: 8 Gaps: 0

US-10-081-935-2 (1-18) x CNS01AF9 (1-660)

Qy 1 MetGlnPheSerHisAlaLeuValAlaAlaGlyLeuAlaSerAla 18

Db 81 ATGCAATTCTCATCTGCCATCATCTCTGCCATCACCGTTGCTCTTCATCGGCC 134

RESULT 7

CNS01AO7

LOCUS Botrytis cinerea strain T4 cDNA library under conditions of

DEFINITION nitrogen deprivation.

ACCESSION AL113183

VERSION AL113183.1 GI:5827802

KEYWORDS cDNA library; nitrogen deprivation.

SOURCE Botryotinia fuckeliana

ORGANISM Botryotinia fuckeliana

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;

AUTHORS Helotiales; Sclerotiniaceae; Botryotinia.

TITLE Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.

JOURNAL Direct Submission

REFERENCE Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,

AUTHORS 78026 Versailles, France

TITLE 2 (bases 1 to 660)

JOURNAL Genoscope.

COMMENT Direct Submission

Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
 CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

The cDNA library to be analyzed within the framework of this
 project was created using a Botrytis cinerea strain which was grown
 under conditions of nitrogen deprivation, which is the normal
 situation for B. cinerea during its development on its host plant.
 The library was produced in an oriented direction, in the pBSII
 vector.

FEATURES

source Location/Qualifiers

1. .660

/organism="Botryotinia fuckeliana"

/mol_type="mRNA"

/strain="T4"

/db_xref="taxon:40559"

/note="Genoscope sequence ID : W59D121"

BASE COUNT 160 a 215 c 132 g 153 t

ORIGIN

Alignment Scores:

Pred. No.: 13.6 Length: 660

Score: 48.00 Matches: 10

Percent Similarity: 72.22% Conservative: 3

Best Local Similarity: 55.56% Mismatches: 5

Query Match: 58.54% Indels: 0

DB: 8 Gaps: 0

US-10-081-935-2 (1-18) x CNS01AO7 (1-660)

Qy 1 MetGlnPheSerHisAlaLeuValAlaAlaGlyLeuAlaSerAla 18

Db 79 ATGCAATTCTCATCTGCCATCATCTCTGCCATCACCGTTGCTCTTCATCGGCC 132

RESULT 8

CNS01B92

LOCUS Botrytis cinerea strain T4 cDNA library under conditions of

DEFINITION nitrogen deprivation.

ACCESSION AL113934

VERSION AL113934.1 GI:5828553

KEYWORDS cDNA library; nitrogen deprivation.

SOURCE Botryotinia fuckeliana

ORGANISM Botryotinia fuckeliana

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;

AUTHORS Helotiales; Sclerotiniaceae; Botryotinia.

TITLE Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.

JOURNAL Direct Submission

REFERENCE Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,

AUTHORS 78026 Versailles, France

TITLE 2 (bases 1 to 660)

JOURNAL Genoscope.

COMMENT Direct Submission

Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
 CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

The cDNA library to be analyzed within the framework of this
 project was created using a Botrytis cinerea strain which was grown
 under conditions of nitrogen deprivation, which is the normal
 situation for B. cinerea during its development on its host plant.
 The library was produced in an oriented direction, in the pBSII
 vector.

FEATURES

source Location/Qualifiers

1. .660

/organism="Botryotinia fuckeliana"

/mol_type="mRNA"

/strain="T4"

/db_xref="taxon:40559"

/note="Genoscope sequence ID : W05H031"

BASE COUNT 161 a 214 c 131 g 154 t

ORIGIN

Alignment Scores:

Pred. No.: 13.6 Length: 660

Score: 48.00 Matches: 10

Percent Similarity: 72.22% Conservative: 3

Best Local Similarity: 55.56% Mismatches: 5

Query Match: 58.54% Indels: 0

DB: 8 Gaps: 0

US-10-081-935-2 (1-18) x CNS01AF9 (1-660)

Qy 1 MetGlnPheSerHisAlaLeuValAlaAlaGlyLeuAlaSerAla 18

Db 81 ATGCAATTCTCATCTGCCATCATCTCTGCCATCACCGTTGCTCTTCATCGGCC 134

CNS01AO7 660 bp mRNA linear PLN 02-SEP-1999
 LOCUS Botrytis cinerea strain T4 cDNA library under conditions of
 DEFINITION nitrogen deprivation.

ACCESSION AL113183

VERSION AL113183.1 GI:5827802

KEYWORDS cDNA library; nitrogen deprivation.

SOURCE Botryotinia fuckeliana

ORGANISM Botryotinia fuckeliana

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;

AUTHORS Helotiales; Sclerotiniaceae; Botryotinia.

TITLE Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.

JOURNAL Direct Submission

REFERENCE Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,

AUTHORS 78026 Versailles, France

TITLE 2 (bases 1 to 660)

JOURNAL Genoscope.

COMMENT Direct Submission

Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
 CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

The cDNA library to be analyzed within the framework of this
 project was created using a Botrytis cinerea strain which was grown
 under conditions of nitrogen deprivation, which is the normal
 situation for B. cinerea during its development on its host plant.
 The library was produced in an oriented direction, in the pBSII
 vector.

FEATURES

source Location/Qualifiers

1. .660

/organism="Botryotinia fuckeliana"

/mol_type="mRNA"

/strain="T4"

/db_xref="taxon:40559"

/note="Genoscope sequence ID : W59D121"

BASE COUNT 160 a 215 c 132 g 153 t

ORIGIN

Alignment Scores:

Pred. No.: 13.6 Length: 660

Score: 48.00 Matches: 10

Percent Similarity: 72.22% Conservative: 3

Best Local Similarity: 55.56% Mismatches: 5

Query Match: 58.54% Indels: 0

DB: 8 Gaps: 0

US-10-081-935-2 (1-18) x CNS01AO7 (1-660)

Qy 1 MetGlnPheSerHisAlaLeuValAlaAlaGlyLeuAlaSerAla 18

Db 79 ATGCAATTCTCATCTGCCATCATCTCTGCCATCACCGTTGCTCTTCATCGGCC 132

RESULT 8

CNS01B92

LOCUS Botrytis cinerea strain T4 cDNA library under conditions of

DEFINITION nitrogen deprivation.

ACCESSION AL113934

VERSION AL113934.1 GI:5828553

KEYWORDS cDNA library; nitrogen deprivation.

SOURCE Botryotinia fuckeliana

ORGANISM Botryotinia fuckeliana

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;

AUTHORS Helotiales; Sclerotiniaceae; Botryotinia.

TITLE Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.

JOURNAL Direct Submission

REFERENCE Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,

AUTHORS 78026 Versailles, France

TITLE 2 (bases 1 to 660)

JOURNAL Genoscope.

COMMENT Direct Submission

Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
 CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

The cDNA library to be analyzed within the framework of this
 project was created using a Botrytis cinerea strain which was grown
 under conditions of nitrogen deprivation, which is the normal
 situation for B. cinerea during its development on its host plant.
 The library was produced in an oriented direction, in the pBSII
 vector.

FEATURES

source Location/Qualifiers

1. .660

/organism="Botryotinia fuckeliana"

/mol_type="mRNA"

/strain="T4"

/db_xref="taxon:40559"

/note="Genoscope sequence ID : W05H031"

BASE COUNT 161 a 214 c 131 g 154 t

ORIGIN

Alignment Scores:

Pred. No.: 13.6 Length: 660

Score: 48.00 Matches: 10

Percent Similarity: 72.22% Conservative: 3

Best Local Similarity: 55.56% Mismatches: 5

Query Match: 58.54% Indels: 0

DB: 8 Gaps: 0

US-10-081-935-2 (1-18) x CNS01AF9 (1-660)

Qy 1 MetGlnPheSerHisAlaLeuValAlaAlaGlyLeuAlaSerAla 18

Db 81 ATGCAATTCTCATCTGCCATCATCTCTGCCATCACCGTTGCTCTTCATCGGCC 134

JOURNAL Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage : CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr)
 COMMENT The cDNA library to be analyzed within the framework of this project was created using a Botrytis cinerea strain which was grown under conditions of nitrogen deprivation, which is the normal situation for B. cinerea during its development on its host plant. The library was produced in an oriented direction, in the pBSII vector.

FEATURES

source Location/Qualifiers
 1..660
 /organism="Botryotinia fuckeliana"
 /mol_type="mRNA"
 /strain="T4"
 /db_xref="taxon:40559"
 /note="Genoscope sequence ID : W21F101"
 BASE COUNT 159 a 203 c 144 g 154 t
 ORIGIN

Alignment Scores:
 Pred. No.: 13.6 Length: 660
 Score: 48.00 Matches: 10
 Percent Similarity: 72.22% Conservative: 3
 Best Local Similarity: 55.56% Mismatches: 5
 Query Match: 58.54% Indels: 0
 DB: 8 Gaps: 0

US-10-081-935-2 (1-18) x CNS01B92 (1-660)

Qy 1 MetGlnPheSerHisAlaLeuValaAlaGlyLeuAlaSerAla 18
 |||||
 Db 39 ATGCAATTCATCTGCCATCATCTCTGCCATCACCCTGCTTGTTCATCGGCC 92

RESULT 9
 CNS01C2S 660 bp mRNA linear PLN 02-SEP-1999
 LOCUS Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation.

ACCESSION AL115804
 VERSION AL115004.1 GI:5829623
 KEYWORDS cDNA library; nitrogen deprivation.
 SOURCE Botryotinia fuckeliana
 ORGANISM Botryotinia fuckeliana
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes; Helotiales; Sclerotiniaceae; Botryotinia.

REFERENCE
 AUTHORS Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr, 78026 Versailles, France
 REFERENCE 2 (bases 1 to 660)
 Genoscope.

Direct Submission
 TITLE Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage : CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 JOURNAL - Web : www.genoscope.cns.fr)

COMMENT The cDNA library to be analyzed within the framework of this project was created using a Botrytis cinerea strain which was grown under conditions of nitrogen deprivation, which is the normal situation for B. cinerea during its development on its host plant. The library was produced in an oriented direction, in the pBSII vector.

FEATURES

source Location/Qualifiers
 1..660
 /organism="Botryotinia fuckeliana"
 /mol_type="mRNA"
 /strain="T4"
 /db_xref="taxon:40559"
 /note="Genoscope sequence ID : W57B021"
 BASE COUNT 158 a 203 c 145 g 154 t
 ORIGIN

Alignment Scores:

Pred. No.: 13.6 Length: 660
 Score: 48.00 Matches: 10
 Percent Similarity: 72.22% Conservative: 3
 Best Local Similarity: 55.56% Mismatches: 5
 Query Match: 58.54% Indels: 0
 DB: 8 Gaps: 0

US-10-081-935-2 (1-18) x CNS01C2S (1-660)

Qy 1 MetGlnPheSerHisAlaLeuValaAlaGlyLeuAlaSerAla 18
 |||||
 Db 38 ATGCAATTCATCTGCCATCATCTCTGCCATCACCCTGCTTGTTCATCGGCC 91

RESULT 10

CNS01CQG 696 bp mRNA linear PLN 03-SEP-1999
 LOCUS Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation.

ACCESSION AL115856
 VERSION AL115856.1 GI:5831072
 KEYWORDS cDNA library; nitrogen deprivation.
 SOURCE Botryotinia fuckeliana

ORGANISM Botryotinia fuckeliana
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes; Helotiales; Sclerotiniaceae; Botryotinia.
 REFERENCE 1 (bases 1 to 696)
 AUTHORS Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr, 78026 Versailles, France

REFERENCE 2 (bases 1 to 696)
 Genoscope.

Direct Submission

TITLE Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage : CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

JOURNAL - Web : www.genoscope.cns.fr)

COMMENT The cDNA library to be analyzed within the framework of this project was created using a Botrytis cinerea strain which was grown under conditions of nitrogen deprivation, which is the normal situation for B. cinerea during its development on its host plant. The library was produced in an oriented direction, in the pBSII vector.

FEATURES Location/Qualifiers

source 1..696
 /organism="Botryotinia fuckeliana"
 /mol_type="mRNA"
 /strain="T4"
 /db_xref="taxon:40559"
 /note="Genoscope sequence ID : W32D051"
 BASE COUNT 174 a 216 c 142 g 164 t
 ORIGIN

Alignment Scores:

Pred. No.: 14.3 Length: 696
 Score: 48.00 Matches: 10
 Percent Similarity: 72.22% Conservative: 3
 Best Local Similarity: 55.56% Mismatches: 5
 Query Match: 58.54% Indels: 0
 DB: 8 Gaps: 0

US-10-081-935-2 (1-18) x CNS01CQG (1-696)

Qy 1 MetGlnPheSerHisAlaLeuValaAlaGlyLeuAlaSerAla 18
 |||||
 Db 80 ATGCAATTCATCTGCCATCATCTCTGCCATCACCCTGCTTGTTCATCGGCC 133

RESULT 11

CNS01DHA 696 bp mRNA linear PLN 03-SEP-1999
 LOCUS Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation.
 ACCESSION AL116822
 VERSION AL116822.1 GI:5832038

KEYWORDS cDNA library; nitrogen deprivation.
 SOURCE Botryotinia fuckeliana
 ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;
 Helotiales; Sclerotiniaceae; Botryotinia.
 REFERENCE 1 (bases 1 to 696)
 AUTHORS Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,
 78026 Versailles, France
 REFERENCE 2 (bases 1 to 696)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
 CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT The cDNA library to be analyzed within the framework of this
 project was created using a Botrytis cinerea strain which was grown
 under conditions of nitrogen deprivation, which is the normal
 situation for B. cinerea during its development on its host plant.
 The library was produced in an oriented direction, in the pBSII
 vector.

FEATURES Location/Qualifiers
 source 1..696
 /organism="Botryotinia fuckeliana"
 /mol_type="mRNA"
 /strain="T4"
 /db_xref="taxon:40559"
 /note="Genoscope sequence ID : W44G081"
 BASE COUNT 175 a 212 c 148 g 161 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 14.3 Length: 696
 Score: 48.00 Matches: 10
 Percent Similarity: 72.22% Conservative: 3
 Best Local Similarity: 55.56% Mismatches: 5
 Query Match: 58.54% Indels: 0
 DB: 8 Gaps: 0
 US-10-081-935-2 (1-18) x CNS01DHA (1-696)

QY 1 MetGlnPheSerHisAlaLeuAlaLeuAlaLeuAlaGlyLeuAlaSerAla 18
 Db 53 ATGCAATTCTCATCTGCCATCATCTCTGCCATCACCGTTGCTTGCATCGGCC 106

RESULT 12
 CNS01BOR 697 bp mRNA linear PLN 02-SEP-1999
 LOCUS Botrytis cinerea strain T4 cDNA library under conditions of
 DEFINITION nitrogen deprivation.
 ACCESSION AL114499
 VERSION AL114499.1 GI:5829118
 KEYWORDS cDNA library; nitrogen deprivation.
 SOURCE Botryotinia fuckeliana
 ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;
 Helotiales; Sclerotiniaceae; Botryotinia.
 REFERENCE 1 (bases 1 to 697)
 AUTHORS Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,
 78026 Versailles, France
 REFERENCE 2 (bases 1 to 697)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
 CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

COMMENT The cDNA library to be analyzed within the framework of this
 project was created using a Botrytis cinerea strain which was grown
 under conditions of nitrogen deprivation, which is the normal
 situation for B. cinerea during its development on its host plant.

The library was produced in an oriented direction, in the pBSII
 vector.
 FEATURES Location/Qualifiers
 source 1..697
 /organism="Botryotinia fuckeliana"
 /mol_type="mRNA"
 /strain="T4"
 /db_xref="taxon:40559"
 /note="Genoscope sequence ID : W15F041"
 BASE COUNT 177 a 219 c 151 g 149 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 14.3 Length: 697
 Score: 48.00 Matches: 10
 Percent Similarity: 72.22% Conservative: 3
 Best Local Similarity: 55.56% Mismatches: 5
 Query Match: 58.54% Indels: 0
 DB: 8 Gaps: 0
 US-10-081-935-2 (1-18) x CNS01BOR (1-697)

QY 1 MetGlnPheSerHisAlaLeuAlaLeuAlaLeuAlaGlyLeuAlaSerAla 18
 Db 59 ATGCAATTCTCATCTGCCATCATCTCTGCCATCACCGTTGCTTGCATCGGCC 112

RESULT 13
 CNS01AOB 720 bp mRNA linear PLN 02-SEP-1999
 LOCUS Botrytis cinerea strain T4 cDNA library under conditions of
 DEFINITION nitrogen deprivation.
 ACCESSION AL113187
 VERSION AL113187.1 GI:5827806
 KEYWORDS cDNA library; nitrogen deprivation.
 SOURCE Botryotinia fuckeliana
 ORGANISM Botryotinia fuckeliana
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;
 Helotiales; Sclerotiniaceae; Botryotinia.
 REFERENCE 1 (bases 1 to 720)
 AUTHORS Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,
 78026 Versailles, France
 REFERENCE 2 (bases 1 to 720)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
 CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

COMMENT The cDNA library to be analyzed within the framework of this
 project was created using a Botrytis cinerea strain which was grown
 under conditions of nitrogen deprivation, which is the normal
 situation for B. cinerea during its development on its host plant.
 The library was produced in an oriented direction, in the pBSII
 vector.

FEATURES Location/Qualifiers
 source 1..720
 /organism="Botryotinia fuckeliana"
 /mol_type="mRNA"
 /strain="T4"
 /db_xref="taxon:40559"
 /note="Genoscope sequence ID : W59D122"
 BASE COUNT 180 a 221 c 145 g 174 t
 ORIGIN

Alignment Scores:
 Pred. No.: 14.8 Length: 720
 Score: 48.00 Matches: 10
 Percent Similarity: 72.22% Conservative: 3
 Best Local Similarity: 55.56% Mismatches: 5
 Query Match: 58.54% Indels: 0
 DB: 8 Gaps: 0


```

Oy 1 MetGlnPheSerHisAlaLeuValaLeuValaAlaGlyLeuAlaSer 17
Db 47008 CTCCAATTTCACAGGATATTACTGACTTCTAGTATCTGCAGGATGGCCAGT 47058

RESULT 16
AC011421 148142 bp DNA linear PRI 18-APR-2000
LOCUS Homo sapiens chromosome 5 clone CTD-230305, complete sequence.
AC011421
AC011421.2 GI:6102634
HTG:
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 148142)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 148142)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 148142)
AUTHORS DOE Joint Genome Center and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1999) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 148142)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (01-FEB-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE 5 (bases 1 to 148142)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE 6 (bases 1 to 148142)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (18-APR-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Oct 23, 1999 this sequence version replaced gi:6013521.
Draft Sequence produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.1.
STS Content:
WI-6744 G05721
WI-6988 G06342.
Location/Qualifiers
1. 148142
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-230305"
BASE COUNT 50576 a 26865 c 24848 g 45853 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 3.31e+03 Length: 148142
Score: 48.00 Matches: 9
Percent Similarity: 82.35% Conservative: 5
Best Local Similarity: 52.94% Mismatches: 3
Query Match: 58.54% Indels: 0
DB: 9 Gaps: 0

```

```

FEATURES
source
Location/Qualifiers
1. 148142
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-230305"

```

```

US-10-081-935-2 (1-18) x AC011421 (1-148142)
Oy 1 MetGlnPheSerHisAlaLeuValaLeuValaAlaGlyLeuAlaSer 17
Db 92942 CTCCAATTTCACAGGATATTACTGACTTCTAGTATCTGCAGGATGGCCAGT 92992

RESULT 17
AC011783 175178 bp DNA linear HTG 13-JUL-2000
LOCUS Homo sapiens clone RP11-16L10, LOW-PASS SEQUENCE SAMPLING.
AC011783
AC011783.3 GI:7144945
HTG: HTGS PHASEO.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 175178)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 175178)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Batna, N., Beckerly, R., Boguslavsky, L., Boukhalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collamore, A.,
Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,
Gallagan, J., Gardy, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Johnson, R., Jones, C., Kann, L., Kardas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
TITLE Direct Submission
JOURNAL Submitted (14-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Mar 3, 2000 this sequence version replaced gi:6446867.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3568
Center clone name: 16_L_10
-----
* NOTE: This record contains 185 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 916: contig of 916 bp in length
* 917 1016: gap of 100 bp
* 1017 1893: contig of 877 bp in length
* 1894 1993: gap of 100 bp
* 1994 2899: contig of 906 bp in length
* 2899 2900: gap of 100 bp
* 2900 3000: contig of 976 bp in length
* 3000 3976: gap of 100 bp
* 3976 5009: contig of 934 bp in length
* 5009 5109: gap of 100 bp

```

5110	6049: contig of 940 bp in length	42919	43018: gap of 100 bp	
6050	6149: gap of 100 bp	43019	43943: contig of 925 bp in length	
6150	7075: contig of 926 bp in length	43944	44043: gap of 100 bp	
7076	7175: gap of 100 bp	44044	44994: contig of 951 bp in length	
8110	8110: contig of 935 bp in length	44995	45094: gap of 100 bp	
8111	8210: gap of 100 bp	45095	46032: contig of 938 bp in length	
8211	9126: contig of 916 bp in length	46033	46132: gap of 100 bp	
9127	9226: gap of 100 bp	46133	47063: contig of 931 bp in length	
9227	10157: contig of 931 bp in length	47064	47163: gap of 100 bp	
10158	10257: gap of 100 bp	47164	48113: contig of 950 bp in length	
10258	11178: contig of 921 bp in length	48114	48213: gap of 100 bp	
11179	11278: gap of 100 bp	48214	49162: contig of 949 bp in length	
11279	12202: contig of 924 bp in length	49163	49262: gap of 100 bp	
12203	13202: gap of 100 bp	49263	50166: contig of 904 bp in length	
13203	13218: contig of 916 bp in length	50167	50266: gap of 100 bp	
13219	13318: gap of 100 bp	50267	51189: contig of 923 bp in length	
13319	14252: contig of 934 bp in length	51190	51289: gap of 100 bp	
14253	14352: gap of 100 bp	51290	52112: contig of 923 bp in length	
14353	15290: contig of 938 bp in length	52113	52213: gap of 100 bp	
15291	15390: gap of 100 bp	52213	53209: contig of 897 bp in length	
15391	16319: contig of 929 bp in length	53210	53309: gap of 100 bp	
16320	16419: gap of 100 bp	53310	54254: contig of 945 bp in length	
16420	17356: contig of 937 bp in length	54255	54354: gap of 100 bp	
17357	17456: gap of 100 bp	54355	55298: contig of 944 bp in length	
17457	18374: contig of 918 bp in length	55299	55398: gap of 100 bp	
18375	18474: gap of 100 bp	55399	56307: contig of 909 bp in length	
18475	19399: contig of 925 bp in length	56308	56407: gap of 100 bp	
19400	19499: gap of 100 bp	56408	57325: contig of 918 bp in length	
19500	20434: contig of 935 bp in length	57326	57425: gap of 100 bp	
20435	20534: gap of 100 bp	57426	58366: contig of 941 bp in length	
20535	21209: contig of 675 bp in length	58367	58466: gap of 100 bp	
21210	21309: gap of 100 bp	58467	59390: contig of 924 bp in length	
21310	22235: contig of 926 bp in length	59391	59490: gap of 100 bp	
22236	22335: gap of 100 bp	59491	60416: contig of 926 bp in length	
22336	22993: contig of 958 bp in length	60417	60516: gap of 100 bp	
22994	23393: gap of 100 bp	60517	61462: contig of 946 bp in length	
23394	24333: contig of 940 bp in length	61463	61562: gap of 100 bp	
24334	24433: gap of 100 bp	61563	62510: contig of 948 bp in length	
24434	25365: contig of 932 bp in length	62511	62610: gap of 100 bp	
25366	25465: gap of 100 bp	62611	63538: contig of 928 bp in length	
25466	26394: contig of 929 bp in length	63539	63638: gap of 100 bp	
26395	26494: gap of 100 bp	63639	64574: contig of 936 bp in length	
26495	27441: contig of 947 bp in length	64575	64674: gap of 100 bp	
27442	27541: gap of 100 bp	64675	65613: contig of 939 bp in length	
27542	28440: contig of 899 bp in length	65614	65713: gap of 100 bp	
28441	28540: gap of 100 bp	65714	66625: contig of 912 bp in length	
28541	29468: contig of 928 bp in length	66626	66725: gap of 100 bp	
29469	29568: gap of 100 bp	66726	67636: contig of 911 bp in length	
29569	30504: contig of 936 bp in length	67637	67736: gap of 100 bp	
30505	30604: gap of 100 bp	67737	68689: contig of 953 bp in length	
30605	31513: contig of 909 bp in length	68690	68789: gap of 100 bp	
31514	31613: gap of 100 bp	68790	69734: contig of 945 bp in length	
31614	32552: contig of 939 bp in length	69735	69834: gap of 100 bp	
32553	32652: gap of 100 bp	69835	70735: contig of 901 bp in length	
32653	33586: contig of 934 bp in length	70736	70835: gap of 100 bp	
33587	33686: gap of 100 bp	70836	71771: contig of 936 bp in length	
33687	34615: contig of 929 bp in length	71772	71871: gap of 100 bp	
34616	34715: gap of 100 bp	71872	72819: contig of 948 bp in length	
34716	35640: contig of 925 bp in length	72820	72919: gap of 100 bp	
35641	35740: gap of 100 bp	72920	73902: contig of 983 bp in length	
35741	36676: contig of 936 bp in length	73903	74002: gap of 100 bp	
36677	36776: gap of 100 bp	74003	74924: contig of 922 bp in length	
36777	37684: contig of 908 bp in length			
37685	37784: gap of 100 bp			
37785	38781: contig of 937 bp in length			
38782	38881: gap of 100 bp			
38882	39829: contig of 948 bp in length			
39829	39929: gap of 100 bp			
39930	40857: contig of 928 bp in length			
40858	40957: gap of 100 bp			
40958	41882: contig of 925 bp in length			
41883	41982: gap of 100 bp			
41983	42918: contig of 936 bp in length			

Alignment Scores:

Pred. No.:	3.93e+03	Length:	175178
Score:	48.00	Matches:	9
Percent Similarity:	82.35%	Conservative:	5
Best Local Similarity:	52.94%	Mismatches:	3
Query Match:	58.54%	Indels:	0
DB:	2	Gaps:	0

US-10-081-935-2 (1-18) x AC011783 (1-175178)

```

QY      1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSer 17
Db      44746 CTCCAATTACAGGATATTACTGACTTTAGTATCGCAGGAATGCCAGT 44696

RESULT 18
LOCUS   GGA252171/c
DEFINITION
Gallus gallus partial mRNA for proprotein convertase PACE4 (PACE4
gene), clone cHPACE.
ACCESSION
AJ252171.1 GI:6706187
VERSION  PACE4 gene; proprotein convertase PACE4.
KEYWORDS
Gallus gallus (Chicken)
SOURCE   Gallus gallus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1 Feldmann,A., Schafer,M.K., Garten,W. and Klenk,H.D.
  Targeted infection of endothelial cells by avian influenza virus
  A/FPV/Rostock/34 (H7N1) in chicken embryos
  J. Virol. 74 (17), 8018-8027 (2000)
20392157
PUBMED  10933711
REFERENCE
2 (bases 1 to 377)
AUTHORS
Feldmann,A.
DIRECT SUBMISSION
SUBMITTED (08-JAN-2000) Feldmann A., Institut fuer Virologie,
Philipps Universitaet Marburg, Robert-Koch-Str. 17, 35037 Marburg,
GERMANY
FEATURES
Source
Location/Qualifiers
1..377
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="cHPACE"
/tissue_type="liver"
/dev_stage="11-day-old"
/country="Germany"
1..377
/gene="PACE4"
<1..>377
/gene="PACE4"
/codon_start=1
/product="proprotein convertase PACE4"
/protein_id="CAB65907.1"
/db_xref="GI:6706188"
/db_xref="SPTREMBL:Q9PT85"
/translation="HKVSHFYGFLGDAEAIIVEAKKTKTPPOHVCVCSLDRVPKYI
RPDHLVRLASTLSACSEOEQHVLYLHVVRSLNAHPRRGDLQISLVSPAGTRSQL
ARRVFDHNSNEGFGKEFWTVHCWG"
BASE COUNT 80 a 107 c 122 g 68 t
ORIGIN

Alignment Scores:
Pred. No.: 11.9 Length: 377
Score: 47.00 Matches: 9
Percent Similarity: 72.2% Conservative: 4
Best Local Similarity: 50.0% Mismatches: 5
Query Match: 57.32% Indels: 0
DB: 5 Gaps: 0

US-10-081-935-2 (1-18) x GGA252171 (1-377)

QY      1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
Db      213 CTCCAGGACAGCAGCTGCTCGCTGCTGCTCGTGCAGGACCTGCTCAGAGT 160

RESULT 19
LOCUS   AE006992
DEFINITION
Mycobacterium tuberculosis CDC1551, section 78 of 280 of the
complete genome.
14483 bp
AE006992
LOCUS   AE006992
DEFINITION
Mycobacterium tuberculosis CDC1551
complete genome.
14483 bp
BCT 27-APR-2001

```

```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS
gene
CDS
gene
CDS

```

```

AE006992 AE000516
AE006992.1 GI:13880691
Mycobacterium tuberculosis CDC1551
Mycobacterium tuberculosis CDC1551
Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;
Corynebacteriaceae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1 (bases 1 to 14483)
Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
Kolony,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
Salzberg,S.L., Delcher,A., Uterback,T., Weidman,J., Khouri,H.,
Gill,J., Mikula,A. and Bishai,W.
Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains
Unpublished
2 (bases 1 to 14483)
Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
Kolony,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
Salzberg,S.L., Delcher,A., Uterback,T., Weidman,J., Khouri,H.,
Gill,J., Mikula,A. and Bishai,W.
Direct Submission
Submitted (25-APR-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1..14483
/organism="Mycobacterium tuberculosis CDC1551"
/mol_type="genomic DNA"
/strain="CDC1551"
/db_xref="taxon:83331"
/note="clinical strain"
/complement(66..560)
/gene="MT111"
/complement(66..560)
/gene="MT111"
/note="similar to SP:P21346; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="transcription elongation factor GreA"
/protein_id="AAK45367.1"
/db_xref="GI:13880692"
/translation="MTDTQVTLTQESHDLKALDQLIANRPVIAEINDRREEGDL
RENGYHAREEQGOEARIQLDLSNAKGEA"KQSGVALPGSVKVVYNGDSD
SETFLIATROEGVSDGKLEVYSPNSPLGALIDAKVGTRTSYTPNGSTVSVTLVSAE
PVHS"
/complement(746..1207)
/gene="MT112"
/complement(746..1207)
/gene="MT112"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK45368.1"
/db_xref="GI:13880693"
/translation="MHPAGVNPQMTHTPIPRDARYGRPLRRRRRVAIALGLVLA
AAGVIAVIGTQRISTASVTLGRLVDDETASVTISVTRSDSPRVACIVRVRA
NGSETGRRELLVPPSEATTQVTTTKSSQPPVADVYCGGTEVPSYLRP"
1282..2148
/gene="MT113"
1282..2148
/gene="MT113"
/note="similar to PID:699274; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="lmbE protein"
/protein_id="AAK45369.1"
/db_xref="GI:13880694"
/translation="MSELRLMVAHVAHPDDESSKGAATLARYADEGHRVLVVTLTGGER

```



```

GEILNPAMDLPDVHGRITAEIRRDGMTKAAEILGVEHTWLGVDGSLPKGDLPPPLPDD
CFARVLELVSTLVRVVRERFPHVMTTYDENGYPHPDHIRCHOUVSAAVEAAGDFC
RFDAGEPTVSLKYLTVHGFRLERMQMLQDEFARHGQGFQEWLAWDPDHDFLTSR
VTRVSCSKYFSQRDDALRAHAHTQIDPNAEFAAPLAWQERLWPTFEFELARSRIIPAR
PPTTELPAGIEP"
2145. .2411
/gene="MT1114"
2145. .2411
/gene="MT1114"
2145. .2411
/gene="MT1114"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK45370.1"
/db_xref="GI:13880695"
/translation="MQILSLVIAEGPGNTGDPFGKASPVGLLVIVLLVIATLFLVR
SMNOQLKKVPKSFDRDHPDLPDQAADEGTRDGPARRPGPPHESG"
2398. .4419
/gene="MT1115"
2398. .4419
/gene="MT1115"
2398. .4419
/gene="MT1115"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK45371.1"
/db_xref="GI:13880696"
/translation="MSPANSGNTLALATSPYLQRHADNPVHQWTPQALAEAAAR
AVPILISVGAHCHWMAHSEFDDDEVAAMNAGFVCIKVDREERPDIDAVMNNAT
VALTGGQWMTGFLTPNGRPFPGCTGYPKAAFLQLLSAISETWRERRAEVQASDHI
AAELRSMASGLPGPEVAPELCDDAVAGVREQDTAHGGFGGAPKPPPSALLLEALMR
HYERTSPAALEAVARTGNMARGGIYDOLGGGFPARYSDVGAMVVPHEKWLVDNAL
LRAYAHWARTGLPARRAAOATFALLDELGSKAPADMTFSSLDADADREGSTVYW
TPVQTEVLGCDGRWAAEVFGTEAGTFEHGTSVLQLPADPDDAARLDRRAALLVA
RLAAQRPARDKVVTSWNGLAITALEASVALDDPDALHAARCAATLRLDLHVVDGR
RRASLGGVDSAAILEDHAMLATGLLALYQLTSEGWLAATGCLLDTAVAHFGDPR
KRWFTDADAEIRLPSDPLDPCATPSGASSIAEALLTAGHVVDGARAEYRWOLAAD
TLRAHAYLLARAPRSAGHMLAVAEAVVRGPLQIYACDLPRSSLADARLAPGGAIV
VGAAGSSALLVGRDRVAGADAAYVCRGVCDLPVTSAAELATALGVPG"
4429. .4548
/gene="MT1116"
4429. .4548
/gene="MT1116"
4429. .4548
/gene="MT1116"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK45372.1"
/db_xref="GI:13880697"
/translation="MAPVHREIRDAVSACRVAIFTLATALGVPCQNTTNHAM"
complement(4518. .5246)
/gene="MT1117"
complement(4518. .5246)
/gene="MT1117"
/notes="similar to PID:887849 GB:U00096 PID:1789266;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="Hly-III family protein"
/protein_id="AAK45373.1"
/db_xref="GI:13880698"
/translation="MSQQADTATTAEARTPAHAHHLVEGVARVLTKPRFRGHIHVS
AGTAVLAGASLVASVAVGSAKAGLTTLAYTAATITMTVTSATYHVRVNNKSATARNWM
KRADHSMIFVFIAGSYTPFALLALPAHDGRVLISVWGGAIGILLKCMKCPAPRSVG
VPLYLLGLGWAVNYTATILHNAGVTALVLLFVGALYSIGGILYAVRWPDWPPTTFGY
HEFFHACTAVAAACHYIAMWVFV"
5357. .6145
/gene="MT1118"
5357. .6145
/gene="MT1118"
/notes="similar to SP:P20182 PID:1196916; identified by
sequence similarity; putative"
/codon_start=1

```

```

/transl_table=11
/product="undecaprenyl diphosphate synthase"
/protein_id="AAK45374.1"
/db_xref="GI:13880699"
/translation="MEIIPRLKEPLYRLYELRLOGLAASKDLPRIHIAVLCDGNRR
WARSAGYDVDSYGYRMAKAIEMLRMCHEAGTELATVYLLSTENLQRPDPELAALIE
ITDVVEEICAPANHVSVRTVGDGLGIEEPARRLCGAVSTPEVASPHVNVAVGYGG
RREIVDAVRALLSKELANGATABELVDVATVEGISENLYTSGQDDPLVIRTSGEORL
SGFLWQSAISEMWFTEAHWPAPRHVDLFLRALRDYSARHRSYGR"
6283. .8591
/gene="MT1118.1"
/notes="similar to GB:AL123456; identified by sequence
similarity; putative"
8768. .9088
/gene="MT1118.2"
8768. .9088
/gene="MT1118.2"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK45375.1"
/db_xref="GI:13880700"
/translation="MPCVGYGDRREFVDVAVEAICENLNTSGOPDPLVIRTSGEOR
LSGHRGPTGVSRRLLRALRDYSTPHASIPYVPPYRSDGIHASRLAVESVFDALAG
RVBL"
9223. .9675
/gene="MT1119"
9223. .9675
/gene="MT1119"
9223. .9675
/gene="MT1119"
/notes="similar to GB:AL123456; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="PE family protein"
/protein_id="AAK45376.1"
/db_xref="GI:13880701"
/translation="MKGKLRMSYMIATPAALTAATAIDIGISAVSVANAAAAAATG
VLAAGDEVLAIAIRLFNANAEYHALSAQVAAPOTLFLVTLTGCGGVFRRRRGRQCV
TAAEHRAAGAGRRQRRRRSGDQWRLRQQRHFHGGQPEFRQHSEHR"
9692. .9859
/gene="MT1120"
9692. .9859
/gene="MT1120"
9692. .9859
/gene="MT1120"
/notes="similar to GB:AL123456; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="PE family protein"
/protein_id="AAK45377.1"
/db_xref="GI:13880702"

```

Alignment Scores:

Pred. No.:	484	Length:	14483
Score:	47.00	Matches:	9
Percent Similarity:	92.86%	Conservative:	4
Best Local Similarity:	64.29%	Mismatches:	1
Query Match:	57.32%	Indels:	0
DB:	1	Gaps:	0

US-10-081-935-2 (1-18) x AE006992 (1-14483)

Qy	3	PhaeH3slaleulleAlaLeuvalAlaAlaGlyLeuhla	16
Db	13148	TTCCGCACGCCCTGGTGCATCGTTGCGCGTGAATTGCGG	13189

RESULT 20

MTV017	LOCUS	67200 bp	DNA	linear	BCT 02-SEP-2002
DEFINITION	Mycobacterium tuberculosis H37Rv complete genome; segment 48/162.				
ACCESSION	AL021897				
VERSION	AL021897.1				
KEYWORDS	GI:3256022				

SOURCE
ORGANISM
Mycobacterium tuberculosis H37Rv
Mycobacterium tuberculosis H37Rv
Bacteria; Actinobacteria; Actinomycetales;
Corynebacteriaceae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE
AUTHORS
1
Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,
Harris, D., Gordon, S.V., Eigmeier, K., Gas, S., Barry III, C.E.,
Tejate, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,
Comor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S.,
Hamlin, N., Holroyd, S., Hornsby, T., Jagsi, K., Krogh, A., McLean, J.,
Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A.,
Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, S.,
Squires, S., Squares, R., Stunton, J.E., Taylor, K., Whitehead, S. and
Barrell, B.G.
Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence
Nature 393 (6685), 537-544 (1998)

JOURNAL
MEDLINE
98295887
PUBMED
9634230
2 (bases 1 to 67200)

REFERENCE
AUTHORS
TITLE
JOURNAL
PARKHILL, J.
Direct Submission
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 26, 1998 this sequence version replaced gi:2896684.
Notes:
Details of M. tuberculosis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have
been renumbered from the original cosmid submissions but the old
gene designations are in brackets after the new gene numbers.
Gene predictions are in brackets after the new gene numbers.
implemented in TBPase (Krogh) supplemented with visual inspection
of positional base preference in codons, especially where there is
CAUTION: In some cases we may not have predicted the correct
initiation codon. Where possible we choose an initiation codon
(atg, gtc, or ttc) which is preceded by an upstream ribosome
binding site sequence (optimally 5-13bp before the initiation
codon). If this cannot be identified we choose the most upstream
initiation codon.

FEATURES
source
Location/Qualifiers
1..67200
/organism="Mycobacterium tuberculosis H37Rv"
/mol_type="genomic DNA"
/strain="H37Rv"
/db_xref="taxon:83332"
66070..>67200
/organism="Mycobacterium tuberculosis H37Rv"
/mol_type="genomic DNA"
/strain="H37Rv"
/db_xref="taxon:83332"
/clone="12268"
<1..66170
/note="fragment designated v017. Does not represent a
physical clone"
178..181
/note="possible RBS, ggaag, for Rv1049"
184..630
/gene="Rv1049"
/gene="Rv1049"
/note="Rv1049, (MTV017.02), len: 148. Probable repressor
similar to many e.g. P74870 (149 aa). FASTA scores:
sp|P74870|P74870NEGATIVE REGULATOR OF EMR LOCUS EMR (149
aa) aov:146 z-score: 200.3 E(): 0.0011; 31.6% identity in
95 aov:146 z-score: 200.3 E(): 0.0011; 31.6% identity in
helix-turn-helix motif at aa 58-79 (Score 1495, +4.28 SD)"
/codon_start=1

gene
CDS
/transl_table=1
/product="hypothetical protein Rv1049"
/protein_id="CA17165.1"
/db_xref="GI:2896684"
/db_xref="SPTREMBL:O53397"
/translation="MGKAFDECACTTTRRAAROLGOAYDRALRPSGLTNTQSTLA
VLSSEGAGIDLTMSLELARIQVETTLNLEVRMDGVVRVAGADARCKRIELT
AKGRAALOKAVPLRQVQAEVTASVGDPRPRRDINLGOAAENCR"
668..672
/note="possible RBS, aggaag, for Rv1050"
679..1584
/gene="Rv1050"
/gene="Rv1050"
679..1584
/note="Rv1050, (MTV017.03), len: 301. Probable
oxidoreductase similar to many e.g. MCV48.22C(141 aa).
FASTA scores: sp|Q10783|Y04M MYCTU PUTATIVE OXIDOREDUCTASE
CY48.22C (141aa) opt:462 z-score: 533.9 E(): 3e-22; 33.6%
identity in265 aa overlap. TBPase score is 0.896"
/codon_start=1
/transl_table=1
/product="hypothetical protein Rv1050"
/protein_id="CA17166.1"
/db_xref="GI:2896687"
/db_xref="SPTREMBL:O53398"
/translation="MARORPRDOVLITGASGIGGPAATKAFAREGAVLAAAREGA
LRVARELAAGCAAWAPLIDVSSSVRAVAVNGEFGIDVFNAGSLVCPVD
AERTLDTRMLELDYGTAVVAREVPLMKQSGSRIMNNSVVGARAFAPAYSS
AMHAIAGPSDLRQELSGSIAVSHPALQTPPLANVADPADPPEPRSLTPVHW
VAAVLDGVARRRARVVPQPLRMVGDFAFSPRYCDRVVRLLESKIFGRIGSRGS
VYRHQPTESAKQAQAQPERGYSAR"
complement (11743..2498)
/gene="Rv1051c"
complement (11743..2498)
/gene="Rv1051c"
/note="Rv1051c, (MTV017.04c), len: 251. Unknown but
similar to protein encoded by lactococcus lactis plasmid
pNP40, L1336837_2 (298 aa). FASTA scores:
U36837|L1336837_2 Lactococcus lactis plasmid pNP40, (298
aa) opt: 194 z-score: 245.5E(): 3.5e-06; 30.3% identity in
155 aa overlap. TBPase score is 0.912. Contains possible
helix-turn-helix motif at aa197-218 (Score 1097, +2.92
SD)"
/codon_start=1
/transl_table=1
/product="hypothetical protein Rv1051c"
/protein_id="CA17167.1"
/db_xref="GI:2896688"
/db_xref="SPTREMBL:O53399"
/translation="MRADVTAEHLTVQVRIAVDIDGVAFNLDTSVQVIRERADY
RGVLRVAVMSGPGWOCIAAMDVSTGEPIAAMPTRVTRDILGEFTILGTAPEITIAE
KGVTLIERGTTSTRMDYVDIVQDRGIDDDLELRARAVAOVRGATLEVAHLAG
YGAIVAQKMAHTEHRCQHWKHPAVGRRNDLDAKQSEMIQVPGVTLRHWRRHS
DIGPASTFLGRVVRDEVSRIKSKESATRR"
3507..3513
/note="possible RBS, aaggtgg, for Rv1052"
3521..3910
/gene="Rv1052"
/gene="Rv1052"
3521..3910
/note="Rv1052, (MTV017.05), len: 129. Unknown"
/codon_start=1
/transl_table=1
/product="hypothetical protein Rv1052"
/protein_id="CA17168.1"
/db_xref="GI:2896689"
/db_xref="SPTREMBL:O53400"
/translation="MDCERGVARHKGISQVGTGCGPRMSOAVSVCASVYRAAYTA
VQVPLRPGVGTPLPHDETALPLVVEVLDKRTITADYDLEQGLQDOVFDLMLPTTA
VEGSLSDILSDHFRDILHAKHFGV"
complement (3809..4084)
/gene="Rv1053c"
complement (3809..4084)

```
/gene="Rv1053c"
/ncbi="Rv1053c" (MTV017.06c), len: 91. Unknown"
/codon_start=1
/transl_table=1
/product="hypothetical protein Rv1053c"
/protein_id="CA17169.1"
/db_xref="GI:2896690"
/db_xref="SPTREMBL:O53401"
/translation="MDSHKVCMNNNTQLPTGPIIGVHFAVDGVRVAYLDLRLCN
TDVEFTSPPPGVPVLYRTKTRVEIADDEMTVEKLKQRAFRNSRHQ"
4726..4932
/gene="Rv1054"
4726..4932
/gene="Rv1054" (MTV017.07), len: 68. Unknown but similar
to hypothetical Mycobacterium tuberculosis protein
MTCY3G12.25 (151 aa shows similarity to integrases) and to
Mycobacterium paratuberculosis integrase MSCI1.1 (191
aa). This orf continues in another frame as MTV017.08 but no
error can be found to account for frameshift. FASTA scores:
279702|MTCY3G12.25 (151 aa) opt: 273 z-score: 364.1 E():
8.8e-13; 64.7% identity in 68 aa overlap; and
L39071|MSCI1.1 (191aa) opt: 105 z-score: 148.5 E(): 0.9;
31.8% identity in 85 aa overlap."
/codon_start=1
/transl_table=1
/product="hypothetical protein Rv1054"
/protein_id="CA17170.1"
/db_xref="GI:2896691"
/db_xref="SPTREMBL:O53402"
/translation="MTGKGVSTTKRDRHVPPEVWRRLHAEPTDPAALVFG
RKGFPLGEYRMFADNADGVGIE"
4935..5171
/gene="Rv1055"
4935..5171
/gene="Rv1055" (MTV017.08), len: 78. Partial orf, first
49aa similar to hypothetical Mycobacterium tuberculosis
protein MTCY3G12.25 (151 aa shows similarity to
integrases) and to Mycobacterium paratuberculosis integrase
L39071|MSCI1.1 (191 aa) and to many other integrases or
transposases. FASTA scores: 279702|MTCY3G12.25 (151 aa) opt:
291 z-score: 428.6 E(): 2.2e-16; 74.3% identity in 70 aa
overlap; and gpl39071|MSCI1.1 (191 aa) opt: 146 z-score:
221.0 E(): 8.3e-05; 52.1% identity in 48 aa overlap"
/codon_start=1
/transl_table=1
/product="hypothetical protein Rv1055"
/protein_id="CA17171.1"

Alignment Scores:
Pred. No.: 2.3e+03 Length: 67200
Score: 47.00 Matches: 9
Percent Similarity: 92.86% Conservative: 4
Best Local Similarity: 64.29% Mismatches: 1
Query Match: 57.32% Indels: 0
Gaps: 0
DB: 1

US-10-081-935-2 (1-18) x MTV017 (1-67200)
Oy 3 pheserhisalaleuilealaleuvalaiaaiaaglyleuvala 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 46872 TTGGCGCACGCCCTGCTGTCATCTGTCGCTGAATTGGC 46913

RESULT 21
AC144792 210652 bp DNA linear HTG 20-MAY-2003
LOCUS AC144792 Mus musculus chromosome UNK clone RP24-solm24, WORKING DRAFT
DEFINITION AC144792.1 GI:30911145
ACCESSION AC144792.1
VERSION AC144792.1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
```

```
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 210652)
AUTHORS
Wilson, R. K.
TITLE
The sequence of Mus musculus clone
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 210652)
AUTHORS
Wilson, R. K.
TITLE
Direct Submission
JOURNAL
Submitted (20-May-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M_BB0501M24
----- Summary Statistics -----
Sequencing vector: M13, 0%
Chemistry: Dye-primer ET; 0% of reads
Assembly: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 209398 bases at least Q40
Consensus quality: 209715 bases at least Q30
Consensus quality: 209918 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 12335: contig of 12335 bp in length
* 12336 12435: gap of unknown length
* 12436 37833: contig of 25398 bp in length
* 37834 37933: gap of unknown length
* 37934 94620: contig of 56687 bp in length
* 94621 94720: gap of unknown length
* 94721 210652: contig of 115932 bp in length.
*
FEATURES
Location/Qualifiers
source
1..210652
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP24-501M24"
1..12335
/note="assembly_name:Contig17"
12436..37833
/note="assembly_name:Contig18"
37934..94620
/note="assembly_name:Contig19"
94721..210652
/note="assembly_name:Contig20"
BASE COUNT 54629 a 48568 c 48742 g 58398 t 315 others
ORIGIN

Alignment Scores:
Pred. No.: 7.35e+03 Length: 210652
Score: 47.00 Matches: 8
Percent Similarity: 82.35% Conservative: 6
Best Local Similarity: 47.06% Mismatches: 3
Query Match: 57.32% Indels: 0
Gaps: 0
DB: 2

US-10-081-935-2 (1-18) x AC144792 (1-210652)
```



```

gene
CDS
    /locus_tag="Mb0915c"
    /complement(6063..6920)
    /locus_tag="Mb0915c"
    /note="Mb0915c", len: 285 aa. Equivalent to Rv0891c,
    len: 285 aa, from Mycobacterium tuberculosis strain H37Rv,
    (99.6% identity in 285 aa overlap). Possible
    transcriptional regulator, highly similar in N-terminus to
    NP_302202.1|NC_002677 possible transcriptional regulator
    from Mycobacterium leprae (1106 aa). Also highly similar
    to several Mycobacterium tuberculosis putative
    transcriptional regulators e.g. Q102|MT02B10.22
    PROBABLE TRANSCRIPTIONAL REGULATOR PROTEIN (1159 aa),
    FASTA scores: opt: 702, E(): 8.3e-40, (50.6% identity in
    247 aa overlap); MT036.21; MT008.44; MT02B10.23. Also
    shows similarity with several adenylate cyclases and
    hydrolases from other organisms."
    /codon_start=1
    /transl_table=1
    /product="POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN"
    /protein_id="CAD93776.1"
    /db_xref="GI:31617667"
    /translation="MLFNVAHNSLPNIDIDHAIIRGDDHPPTCAKCVAGRISALGS
    LDIRYHSRCYAAPDYRCFVPPRRRVLIANGLDVSRIPPTGTTLLADVEEST
    HLMQCPEDWATRAIADHTVESEITNMGVOPVRYEGSFVAATRASDAACALD
    LQRTSLAPRLRLGLHTGEVLRLDELIVPPIINTARLRDLAHGQVLSATDLYLT
    GRLPADAMLVLDLGRHPLRGLPPEVMVQLCPDIRREKPEPLRTAKSSPTSLPQFPTT
    FVGRRAQIS"
    7318..8805
    /locus_tag="Mb0916"
    7318..8805
    /locus_tag="Mb0916"
    /EC_number="1.14.-.-"
    /note="Mb0916", len: 495 aa. Equivalent to Rv0892, len:
    495 aa, from Mycobacterium tuberculosis strain H37Rv,
    (100.0% identity in 495 aa overlap). Probable
    monooxygenase (EC 1.14.-.-), highly similar to others e.g.
    NP_250787.1|NC_002516 probable flavin-binding
    monooxygenase from Pseudomonas aeruginosa (491 aa);
    CAB59668.1|AL132674 monooxygenase from Streptomyces
    coelicolor (519 aa); P12015|CYMO ACIS cyclohexanone
    monooxygenase from Acinetobacter sp. (542 aa), FASTA
    scores: opt: 489, E(): 6.8e-26, (30.3% identity in 492 aa
    overlap); etc. Also highly similar to Rv0565c, Rv3854c,
    hydrophobic stretch at N-terminus."
    /codon_start=1
    /transl_table=1
    /product="PROBABLE MONOOXYGENASE"
    /protein_id="CAD93777.1"
    /db_xref="GI:31617668"
    /translation="MTGRCPYAVAVGAGMSGMCVAILLSAGITDVCIEKADVGST
    WRNTYVGLTCDVPSRLYOXFAKNPMWTFSGGSIQYLRIGARVGLRHIRRG
    ATVVSARFDDGRWLRTDSCSTESTVFLIATGLVHHPRLPTAGLDDPGTYFHSAR
    WDRHTVPLIGRIAVICGTSGTQVUCLVAGLAVGAKVTKFQTAQVLPWPPIRSKLR
    VFRARFPCGSLAVKAYSLSFETFAVALSNPGLHRKLVGAVCRASLRVSDPRLRL
    TPDIPEWCKRLVMSGFGYRAIQRDVGLVAGIDVHEHGIIVTDGVLHSDVIVLAT
    GFDSHAFFRPMOLTGDRIRIDVMDGPPAHQVAPGPFNFMMGLPSPVGNFPL
    TAAVESQAEHIIVQIKRMHGEFDTMEPKSAATEAVYTVLRAMPVNTVMTGDSWYL

```

Alignment Scores:

```

Pred. No.: 1.15e+04 Length: 327650
Score: 47.00 Matches: 9
Percent Similarity: 92.86% Conservative: 4
Best Local Similarity: 64.29% Mismatches: 1
Query Match: 57.32% Indels: 0
DB: 1 Gaps: 0

```

US-10-081-935-2 (1-18) x BX24837 (1-327650)

OY 3 PheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAla 16

```

Db 233446 TTCCGACACGCCCTGTGTGATCGATGTCCTCAATTGGCC 233487
RESULT 23
G32057 1204 bp DNA linear STS 27-OCT-1998
LOCUS MCM301 Chicken RPMACrooijmans Gallus gallus STS genomic, sequence
tagged site.
DEFINITION
ACCESSION G32057
VERSION G32057.1 GI:2138003
KEYWORDS STS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 1204)
REFERENCE Crooijmans R.P., Dijkhof R.J., van der Poel J.J. and Groenen M.A.
New microsatellite markers in chicken optimized for automated
fluorescent genotyping
JOURNAL Anim. Genet. 28 (6), 427-437 (1997)
MEDLINE 98251489
PUBMED 9589584
COMMENT
Contact: Richard Crooijmans
Department of Animal Breeding
Wageningen Agricultural University
Box 338, 6700 AH Wageningen, The Netherlands
Email: richard.crooijmans@alg.vf.war.nl
Primer A: GGAGGAGACCACTGATTC
Primer B: AGGCTAGACGTACCAAGTGC
STS size: 278
PCR Profile:
  Presoak: 94 degrees C for 5 minute (s) :
  Denaturation: 94 degrees C for 30 seconds
  Annealing: 55 degrees C for 45 seconds
  Polymerization: 72 degrees C for 90 seconds
  PCR Cycles: 35
  Thermal Cycler: MJ research
Protocol:
  Template: 10-100 ng
  Primer: each 295 nM
  dNTPs: each 200 uM
  Polymerase Goldstar: 0.4 U
  Total Vol: 12 kl
Buffer:
  MgCl2: 1.5 mM
  KCl: 50 mM
  Tris-HCl: 10 mM
  Tetramethylammoniumchloride (TMAc): 1 mM
  pH: 8.3

```

FEATURES

size range: 266-284.
 Location/Qualifiers
 1..1204
 /organism="Gallus gallus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9031"
 /clone_lib="Chicken RPMACrooijmans"

STS
 primer_bind 290..567
 primer_bind 290..310
 BASE COUNT 260 a 279 c 294 g 360 t 11 others
 ORIGIN

Alignment Scores:

```

Pred. No.: 60 Length: 1204
Score: 46.00 Matches: 9
Percent Similarity: 92.31% Conservative: 3
Best Local Similarity: 69.23% Mismatches: 1
Query Match: 56.10% Indels: 0
DB: 11 Gaps: 0

```


Score: 46.00 Matches: 10
Percent Similarity: 84.62% Conservative: 1
Best Local Similarity: 76.92% Mismatches: 2
Query Match: 56.10% Indels: 0
DB: 8 Gaps: 0
US-10-081-935-2 (1-18) x AY036106 (1-4029)
QY 5 HisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSer 17
Db 2487 CACGACATTGAGCCCTTGTTGCTGTGACTGACGACGCC 2449
RESULT 26
LOCUS MMU307017 8117 bp mRNA linear ROD 14-MAY-2001
DEFINITION Mus musculus mRNA for putative ubiquitin-specific protease (Uspsy
gene)
ACCESSION AJ307017 GI:14041781
VERSION AJ307017.1
KEYWORDS ubiquitin-specific protease; Uspsy gene.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Hall, N.M.
JOURNAL Thesis (2001) Department of Pathology, University of Cambridge,
Cambridge, United Kingdom
REFERENCE 2 (bases 1 to 8117)
AUTHORS Hall, N.M.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2001) Hall N.M., Department of Pathology,
University of Cambridge, Tennis Court Road, Cambridge, CB2 1QP,
UNITED KINGDOM
FEATURES
source
1..8117
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/chromosome="Y"
/map="Xrb"
/sex="Male"
/tissue_type="testis"
1..8117
/gene="Uspsy"
/gene="Uspsy"
419..8089
/gene="Uspsy"
/codon_start=1
/product="putative ubiquitin-specific protease"
/protein_id="CAC3831.1"
/db_xref="GI:14041782"
/translation="MTITKSGPVGENSGQCTSDGQPOPSFOONQISSSDSNETSP
TTPYEQOGAPQHEEDSPFPTDPAKLEMDNRSRWVVPVLPKLELVLETISI
DLTKGLDVKSEACQPRFDVLTYSFKILMEADVAAGFEIHRCIINNTIRLVLCV
AKLSODMPFLLEIALNPCKFHVNGARCESVSSVSPDEBELFACSPDLSPK
GMVLDLINTFTGLNGFOILHDFRTSGALNTOIAPKPGQCEPFTOTLRKFI
PVEIVPQILKLTNEBELKERTKYADDTISMTKFLANLSRIPGGEVTKNLEFI
RLKTLRLQIISFNGKNALEINIKLVSSVYTHGSEEWLVTERTWELQON
NILSTIVLQDSLHQPOYVERKIEKILRFVKEKALTQIDNNTIWAQAGHEAIVKXVD
LLAKLAWNSPEQDLHDFCFKASWTNASKOREKLELIRLLEDDXGVAAHVLD
LLMNLASDDVDVIMDLASAKIKILDYSGSDRDTOKIOMIDCFIEPRTNKVI
PALKOIKEIGLFEAPONLSOTHSPPVFRHDLISQHNHALVTVAENLAAYNA
SIRLYARHEDVDPTVILGRSHYVQVOERLNTLRLLKDGOLMLCVSOAKIOWC
LAENAVYSDBEACFMYSKLMGDEPDLHPDINKEFFESNVQLDPSLLENGMKCFE
RFFTVNCREKMLTKRKTYMDDLDLGLYLMKVVYIQSNDIISRAIDILKEIYTS
LGRPLQANQVYIHEDFIQSCFDRLKASVDTLCVLDSEKDNIFSCARQEAIMVRLTV
LREYISEYSDSYHEERMLPMSRAFGKHLSTVRFPHQGEVEDLILSHNTATGS
VRRCILNNVNVVAHTKIELFELGELVASEDDRLKEQLNKXSLITAKFIQINSNM
PSPDSSSDSAGPRSHNNYRVSNPMEKICPGYMSIQSPRIYSIFLQVADLGS
MLTVPTLADGARILMKLMPDSTTEQLRALSCHDVNIGERLQSLSLTFGSSASQ
VLYLTVYTTILMPAGAPLADISDFQYHPLKSGGLPLVLMLIQNNFLPNTDETNR
DAVFSALKIATLLITVIGYHQAIAEACVPADGTDKTPINQVTHDOAVVLQNALQ

SIRPSSRCMLRNVSVHLAOLQISGLASRYIDICVIRAIQKILMAAGCSLELVFSPN
EDITETTYMTTSTNSVLEKDEQVCCLELVMTTLCFALIPTAMDSLNKEKANSQVYID
LLLCPSKTVRQLOAQEOFLCTCRCCMKHRPLPFTLTLLFTLLGGAANEKGRSDVYF
TLRLRLLYAVANSNIQVNPVVDLNDLDEIDMKRVADYIKNTGETVNGHGLVGT
KELLSFQSPKKYHIGCKTGGANLVKELIDYFIFPASKAYIQYMGSELPICQALPVC
GSPATINAGFELVALAFCGVNRKQIKVNCULTELYIGTPTTCVAGMEWELPVPVG
RPRKFGVNLKAGATCYMNSVYIQQLFMTPSIRNSLAIADISMDPTDDIIFKXKQDE
NNVDREDVFRYPHOFEDKPTLSKEDREKYNIALVRLQITFGLAASOLOYYIPKG
FWQGFRLGEPVNLREOHDALEFVNSLVDSLDEAKALGYPTVLSKVLGGSFADOKIC
QGCHERYCEESFTLVNDVIRNHQNLDSLEOYVGDLLLEGANAHYCKDCDKQVTVK
RLILKPLSVLTIOQLKREDYMERECALKFNDYFEFFPRELMEPYTVAGATLKGDSV
NPOTOLIKONQSESVIPGSTKYRLVGVLYHSGANGGHVSYIIORNGKDSKSHWF
KFDDEGDETCMDDDDEEMKNOCCFGEYVGEVFDHMKRMSYRQKRMNAYILFERM
DITDEDEITTYISELFTFRPHQINSPIENSVMKONQVPHNQPSLEYTFOPIKLL
LTCNAVYLSAPAGODHLLEPAEDITMISIOLASRFLFTTGHTKKIIRGPAWMDAL
CILRHSKNVAFWYVNVLFVNSRNFSELYLECPAEIRGFAXLIVFIAHPSLDQDS
SPSPFTSPFANRPGYSOIYDNLISLSDLLKLVLSLREVSCHGHLQOYFPLFYMA
SLGLAEKTOILKLVVPATFPMVLSIDEGCPVVKOYALSLKHSVSOILIRCCSVSR
MOSSINGNPLPVPFQDNLSQPIPTIQNNADILFMTTTHKKYIEDCSNSEDYVK
LIFCCWENPQFSCVLSLELMQVASHAYELQPYLDLQIILFEDSOAHRILHAKL
GIPNDQDLFDITIGHSKNHQKRAVQCIKMVTLFNSCPVAYOILQDNGDLNKKMTWA
MEWLGDLELRKPYSGNPQYTSNMSPVQSNETANGYFLKSHSAKMLTKACDLYPE
EDPDQDLADEVSHAPQDRTRTFVLYSHRSYVQONVVPQPSGPASHHLLNPNQKNDKP
OETHSNEISISCLIKQ"
BASE COUNT 2634 a 1481 c 1586 g 2416 t
ORIGIN
Alignment Scores:
Pred. No.: 417 Length: 8117
Score: 46.00 Matches: 9
Percent Similarity: 76.47% Conservative: 4
Best Local Similarity: 52.94% Mismatches: 4
Query Match: 56.10% Indels: 0
DB: 10 Gaps: 0
US-10-081-935-2 (1-18) x MMU307017 (1-8117)
QY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSer 17
Db 2231 CTTGACATTAATCATGCTTTAGTTACTTGTAAGCAGAAACCTTGACGACT 2281
RESULT 27
LOCUS AX305688 10323 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 439 from Patent WO01818188.
ACCESSION AX305688
VERSION AX305688.1 GI:17645121
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T. and Ishii, Y.
JOURNAL Method for examining ischemic conditions
TITLE Patent: WO 01818188-A 439 22-NOV-2001.
JOURNAL School Juridical Person Nihon University (JP)
FEATURES
source
1..10323
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 3055 a 1989 c 2201 g 3078 t
ORIGIN
Alignment Scores:
Pred. No.: 532 Length: 10323
Score: 46.00 Matches: 9
Percent Similarity: 76.47% Conservative: 4
Best Local Similarity: 52.94% Mismatches: 4
Query Match: 56.10% Indels: 0
DB: 6 Gaps: 0

University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Apr 3, 2003 this sequence version replaced gi:29373197.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: H_AA087252A06

Location/Qualifiers
1. 26013
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/clone="XXFOS-87252A6"

BASE COUNT 6788 a 5982 c 5844 g 7399 t
ORIGIN

Alignment Scores:
Pred. No.: 1.36e+03 Length: 26013
Score: 46.00 Matches: 11
Percent Similarity: 78.57% Conservative: 0
Best Local Similarity: 78.57% Mismatches: 3
Query Match: 56.10% Indels: 0
DB: 9 Gaps: 0

US-10-081-935-2 (1-18) x AC142230 (1-26013)

OY 2 GlnPhSeHsHAlaLeuHleAlaLeuValAlaAGlyLeuAlaSer 15
|||||
Db 5372 CAGATTCTCCACGCCCTTCCAGCCTTCTTCCAGCAGCGCCCTT 5413

RESULT 30 AC138990 125439 bp DNA linear PRI 27-MAR-2003
LOCUS Homo sapiens chromosome 5 clone RP13-708P15, complete sequence.
AC138990
AC138990.2 GI:29294040
VERSION HTG.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 125439)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submision
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 125439)
REFERENCE 2 DOE Joint Genome Institute.
JOURNAL Direct Submision
AUTHORS Submitted (22-JAN-2003) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94596, USA
3 (bases 1 to 125439)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submision
JOURNAL Submitted (27-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Mar 27, 2003 this sequence version replaced gi:27819484.
Drat Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.hgsc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
NOTE: This sequence is not the entire sequence of the clone
(entire sequence is 153.5kb). It is clipped at the overlap with
AC114316. The number of basepairs overlapped is 118195.
Location/Qualifiers
1. 125439

FEATURES
source

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP13-708P15"
BASE COUNT 40037 a 22627 c 22407 g 40368 t
ORIGIN

Alignment Scores:
Pred. No.: 6.73e+03 Length: 125439
Score: 46.00 Matches: 9
Percent Similarity: 81.25% Conservative: 4
Best Local Similarity: 56.25% Mismatches: 3
Query Match: 56.10% Indels: 0
DB: 9 Gaps: 0

US-10-081-935-2 (1-18) x AC138990 (1-125439)

OY 2 GlnPhSeHsHAlaLeuHleAlaLeuValAlaAGlyLeuAlaSer 17
|||||
Db 43208 AATACCATCATGCGCTATTAGCAAAAGTGCGAGTGGTTATCTCA 43255

RESULT 31 AC073440 133445 bp DNA linear HTG 08-JAN-2003
LOCUS Homo sapiens chromosome 12 clone RP11-100E14, WORKING DRAFT
DEFINITION
SEQUENCE, 4 unordered pieces.
AC073440
AC073440.12 GI:20335621
VERSION HTG: HTGS PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 133445)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barberia,J., Benton,J., Binsge,K., Binkenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyte,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroli,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Doutwaite,K.U., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsf.,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsosn,E., Kelly,S., Khan,U., King,L., Korvach,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,U., Li,Z., Lichtege,O., Lieu,C., Liu,J., Liu,W., Louleaged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mahwinny,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okwunu,G.,
Oragunye,N., Oyaleo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojuboakan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshbari,N., Stason,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Sveteck,A., Tabot,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tamey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Uemami,K., Vasquez,L., Vera,V., Villalton,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K.,

Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 133445)
Mortley, K. C.
Direct Submission
Submitted (18-JUN-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 133445)
Mortley, K. C.
Direct Submission
Submitted (08-JAN-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Apr 28, 2002 this sequence version replaced gi:16117941.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Drafting Center Code: BCM
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HBCF
Center clone name: RP11-100E14
----- Summary Statistics
Sequencing vector: MJ3
Chemistry: Dye-terminator Big Dye 811 of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 128546 bases at least Q40
Consensus quality: 130561 bases at least Q30
Consensus quality: 131661 bases at least Q20
Estimated insert size: 134201; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2013: contig of 2013 bp in length
2 2014 2113: gap of unknown length
3 2114 31789: contig of 29676 bp in length
4 31790 31889: gap of unknown length
5 31890 79584: contig of 47695 bp in length
6 79585 79684: gap of unknown length
7 79685 133445: contig of 53761 bp in length.

FEATURES
source
1..133445
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-100E14"
BASE COUNT 41186 a 25768 c 25551 g 40640 t 300 others
ORIGIN
Alignment Scores:
Pred. No.: 7.16e+03 Length: 133445
Score: 46.00 Matches: 7
Percent Similarity: 93.75% Conservative: 8
Best Local Similarity: 43.75% Mismatches: 1
Query Match: 56.10% Indels: 0
DB: 2 Gaps: 0
US-10-081-935-2 (1-18) x AC073440 (1-133445)

Cy 1 MetClnPhseerHsAlaLeu1eAlaLeuValAlaAclgyleuAla 16
Db 117092 TTGACGTTTACTCAGCTCTCTCCCTTCCTGACAGCTGAAAGCCA 117045
RESULT 32
AC138980
LOCUS
DEFINITION
Homo sapiens chromosome 5 clone RP13-169F15, WORKING DRAFT
SEQUENCE. 2 unordered pieces.
AC138980 147008 bp DNA linear HTG 22-JAN-2003
AC138980
AC138980.1 GI:27819474
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS
DOE Joint Genome Institute.
TITLE
JOURNAL
Unpublished
2 (bases 1 to 147008)
Sequencing of Human Chromosome 5
DOE Joint Genome Institute.
Direct Submission
Submitted (22-JAN-2003) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
----- Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 3069541
Center clone name: RPCT-13_169F15

Summary Statistics
Consensus quality: 146757 bases at least Q40
Consensus quality: 146840 bases at least Q30
Consensus quality: 146855 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Estimated insert size: 146908; sum-of-contigs estimation
Quality coverage: 15.57 in Q20 bases; agarose-fp estimation
Quality coverage: 18.54 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 28773: contig of 28773 bp in length
2 28774 28874: gap of unknown length
3 28874 147008: contig of 118135 bp in length.

FEATURES
source
1..147008
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP13-169F15"
/clone_1fb="RPCT human BAC library 13"
BASE COUNT 48290 a 26527 c 25955 g 46136 t 100 others
ORIGIN
Alignment Scores:
Pred. No.: 7.9e+03 Length: 147008
Score: 46.00 Matches: 9
Percent Similarity: 81.25% Conservative: 4
Best Local Similarity: 56.25% Mismatches: 3
Query Match: 56.10% Indels: 0
DB: 2 Gaps: 0

protein, the 5' end of the USP9X gene for ubiquitin specific protease 9 (Drosophila fat facets related) and a Cpg island, complete sequence.

ACCESSION

AL391259

VERSION

AL391259.15 GI:11322864

KEYWORDS

HTG; chloride; Cpg island; intracellular; protease; ubiquitin; USP9X.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 163520)

AUTHORS

Whitehead,S.

JOURNAL

Submitted (06-DEC-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequests@sanger.ac.uk

COMMENT

On Nov 23, 2000 this sequence version replaced gi11322140.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Emi, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/ChrX>

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-469E19 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see

<http://www.choi.org/bacpac/home.htm>

VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-469E19. It may be shorter because we sequence overlapping

sections only once, except for a short overlap.

The true left end of clone RP11-469E19 is at 1 in this sequence. The true left end of clone RP5-1051C18 is at 185 in this sequence. The true left end of clone RP5-1172N10 is at 163421 in this sequence.

FEATURES

Location/Qualifiers

source

1..163520

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="X"

/map="p11.4-21.2"

/clone="RP11-469E19"

/clone.lib="RPCI-11.2"

repeat_region

408..615

/note="AluJo repeat: matches 87. .286 of consensus"

repeat_region

1462..1669

/note="MLTID repeat: matches 217. .461 of consensus"

repeat_region

1726..1897

/note="FRAM repeat: matches 0. .173 of consensus"

repeat_region

1923..2107

/note="MLTID repeat: matches 3. .184 of consensus"

repeat_region

2147..2448

/note="AluSc repeat: matches 1. .302 of consensus"

repeat_region

2490..2647

/note="AluSc repeat: matches 139. .296 of consensus"

repeat_region

2678..2970

misc_feature

repeat_region

misc_feature

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

/note="AluJo repeat: matches 17. .308 of consensus"

/note="match(2970..3296)

/note="match: GSS: Em:AQ279697"

2983..3228

/note="AluJb repeat: matches 24. .293 of consensus"

/note="match: GSS: Em:AQ278728"

4145..4279

/note="FLAM_C repeat: matches 12. .143 of consensus"

4693..4907

/note="AluJb repeat: matches 85. .309 of consensus"

4918..5199

/note="AluSc repeat: matches 1. .282 of consensus"

5200..5275

/note="38 copies 2 mer ta 97% conserved"

5570..5872

/note="AluSc repeat: matches 1. .303 of consensus"

6323..6576

/note="AluJo repeat: matches 40. .281 of consensus"

6627..6702

/note="LTRJ3 repeat: matches 2. .84 of consensus"

6761..6931

/note="MIR repeat: matches 14. .191 of consensus"

7009..7140

/note="AluJo repeat: matches 1. .136 of consensus"

7141..7435

/note="AluJb repeat: matches 1. .296 of consensus"

7436..7625

/note="AluJo repeat: matches 136. .312 of consensus"

7626..7733

/note="54 copies 2 mer aa 77% conserved"

8520..8657

/note="MIR repeat: matches 3. .147 of consensus"

8850..8977

/note="AluSc repeat: matches 9. .137 of consensus"

8978..9298

/note="AluJg repeat: matches 1. .308 of consensus"

9299..9451

/note="AluSc repeat: matches 137. .286 of consensus"

9567..9919

/note="FLMCS repeat: matches 7156. .7496 of consensus"

9924..10086

/note="Char11el repeat: matches 2593. .2756 of consensus"

10089..10233

/note="FRAM repeat: matches 2. .146 of consensus"

10256..10557

/note="AluSc repeat: matches 3. .304 of consensus"

10588..10946

/note="Char11el repeat: matches 2161. .2574 of consensus"

10945..11034

/note="Char11el repeat: matches 1287. .1379 of consensus"

11039..11322

/note="AluSc repeat: matches 2. .302 of consensus"

11333..11367

/note="U2 repeat: matches 1. .35 of consensus"

11381..11519

/note="FRAM repeat: matches 2. .146 of consensus"

11630..12322

/note="Char11el repeat: matches 381. .1104 of consensus"

12345..12460

/note="FLAM_A repeat: matches 1. .115 of consensus"

12480..12804

/note="Char11el repeat: matches 33. .367 of consensus"

12896..13033

/note="FLMCS repeat: matches 5662. .5791 of consensus"

13097..13178

/note="FLAM_A repeat: matches 27. .109 of consensus"

13184..13491

/note="FLMCS repeat: matches 5323. .5613 of consensus"

13500..13636

/note="LTR41 repeat: matches 13. .148 of consensus"

13637..13940

/note="AluSc repeat: matches 1. .305 of consensus"

```

repeat_region      13941..14089
                    /note="LTR41 repeat: matches 148..301 of consensus"
repeat_region      14112..14144
                    /note="LTR41 repeat: matches 418..453 of consensus"
repeat_region      14145..14452
                    /note="ALUdo repeat: matches 1..305 of consensus"
repeat_region      14852..15028
                    /note="L1MD1 repeat: matches 5705..5884 of consensus"
repeat_region      15127..15382
                    /note="LTR23 repeat: matches 204..436 of consensus"
repeat_region      15383..15480
                    /note="BUR1 repeat: matches 6167..6271 of consensus"
repeat_region      15482..15563
                    /note="LTR23 repeat: matches 36..109 of consensus"
repeat_region      15595..15668
                    /note="L1P4 repeat: matches 5258..5451 of consensus"
repeat_region      15980..16075
                    /note="MERSA repeat: matches 13..112 of consensus"
repeat_region      16796..17092
                    /note="ALUSP repeat: matches 1..296 of consensus"
repeat_region      17479..17743
                    /note="ALUSx repeat: matches 1..266 of consensus"
repeat_region      17743..17987
                    /note="ALUy repeat: matches 46..290 of consensus"
repeat_region      18468..18553
                    /note="MLT1H repeat: matches 93..172 of consensus"
repeat_region      18606..18832
                    /note="ALUSg/x repeat: matches 70..296 of consensus"
repeat_region      18833..18998
                    /note="ALUSx repeat: matches 1..166 of consensus"
repeat_region      19100..19227
                    /note="L1PA11 repeat: matches 5718..5849 of consensus"
repeat_region      19346..19646
                    /note="ALUSg repeat: matches 1..301 of consensus"
repeat_region      19647..19730
                    /note="42 copies 2 mer ta 65% conserved"
repeat_region      19835..20139
                    /note="ALUSx repeat: matches 3..297 of consensus"
repeat_region      20954..21192
                    /note="ALUub repeat: matches 77..311 of consensus"
repeat_region      21361..21479
                    /note="ALUSg/x repeat: matches 1..120 of consensus"
repeat_region      21492..21791
                    /note="ALUSg repeat: matches 1..293 of consensus"
repeat_region      21922..22076
                    /note="MERSA repeat: matches 26..189 of consensus"
repeat_region      22091..22191
                    /note="MIR repeat: matches 33..134 of consensus"
repeat_region      22249..22928
                    /note="MER51B repeat: matches 1..617 of consensus"
repeat_region      23550..23861

Alignment Scores:
Pred. No.:      8.8e+03      Length:      163520
Score:          46.00      Matches:      9
Percent Similarity: 76.47%      Conservative: 4
Best Local Similarity: 52.94%      Mismatches: 4
Query Match:      56.10%      Indels:      0
DB:              9      Gaps:      0

US-10-081-935-2 (1-18) x ALJ91259 (1-163520)
Oy      1      MecGlnPheserHisAlaLeuValAlaIaGlyLeuAlaser 17
Db      142090      CTTCAACACATCATGCCCTAGTACTTGCTAGCAGAAAACCTTGCACT 142140

RESULT 36
ALJ56864/c      ALJ56864      166876 bp      DNA      linear      HTG 13-JUN-2001
LOCUS
DEFINITION      Homo sapiens chromosome 1 clone RP5-1051C18, *** SEQUENCING IN
PROGRESS ***, 26 unordered pieces.
ACCESSION      ALJ56864
VERSION      ALJ56864.6      GI:9926631

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

```

HTG: HTGS PHASE1; HTGS_CANCELLED.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Plumb.B.
Direct Submission
Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: clonerequests@sanger.ac.uk
On Aug 25, 2000 this sequence version replaced gi:9214009.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: dj1051C18
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 151580 bases at least Q40
Consensus quality: 156291 bases at least Q30
Consensus quality: 159609 bases at least Q20
Insert size: 164376; sum-of-contigs
Insert size: 184764; 6.6% error; agarose-fp
Quality coverage: 2.91x in Q20 bases; sum-of-contigs
Quality coverage: 2.72x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
7483: contig of 7483 bp in length
7484      7583: gap of 100 bp
7584      1184: contig of 3601 bp in length
1185      1184: gap of 100 bp
1185      1874: contig of 7458 bp in length
1874      1874: gap of 100 bp
1874      2274: contig of 3902 bp in length
2274      2284: gap of 100 bp
2284      3019: contig of 7347 bp in length
3019      3029: gap of 100 bp
3029      3695: contig of 6660 bp in length
3695      3705: gap of 100 bp
3705      4110: contig of 4050 bp in length
4110      4120: gap of 100 bp
4120      4415: contig of 2950 bp in length
4415      4425: gap of 100 bp
4425      4810: contig of 3857 bp in length
4810      4820: gap of 100 bp
4820      4820: contig of 10206 bp in length
4820      5841: gap of 100 bp
5841      5851: gap of 100 bp
5851      7724: contig of 18735 bp in length
7724      7734: gap of 100 bp
7734      8265: contig of 5307 bp in length
8265      8276: gap of 100 bp
8276      8545: contig of 2696 bp in length
8545      8552: gap of 100 bp
8552      9037: contig of 4826 bp in length
9037      9047: gap of 100 bp
9047      9547: contig of 4995 bp in length
9547      9557: gap of 100 bp
9557      9853: contig of 2980 bp in length
9853      9865: gap of 100 bp
9865      10535: contig of 6882 bp in length
10535      10536: gap of 100 bp

```

```
* 105636 114741: contig of 9106 bp in length
* 114742 114841: gap of 100 bp
* 114842 118544: contig of 3703 bp in length
* 118545 118644: gap of 100 bp
* 118645 130337: contig of 11693 bp in length
* 130338 130437: gap of 100 bp
* 130438 144196: contig of 13759 bp in length
* 144197 144296: gap of 100 bp
* 144297 149313: contig of 5017 bp in length
* 149314 149413: gap of 100 bp
* 149414 152428: contig of 3015 bp in length
* 152429 152528: gap of 100 bp
* 154910 154910: contig of 2382 bp in length
* 154911 155010: gap of 100 bp
* 155011 163649: contig of 8639 bp in length
* 163650 163750: gap of 100 bp
* 163750 166876: contig of 3127 bp in length.

FEATURES
SOURCE
1..166876
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP5-1051C18"
/clone_lib="RPCI-5"

1..7483
/note="assembly_fragment:01007
clone_end:77
vector_side:left"
7584..11184
/note="assembly_fragment:00504
fragment_chain:1"
11285..18742
/note="assembly_fragment:00626
fragment_chain:1"
18843..22744
/note="assembly_fragment:00511
fragment_chain:2"
22845..30191
/note="assembly_fragment:01426
fragment_chain:2"
30292..36951
/note="assembly_fragment:00759
fragment_chain:3"
37052..41101
/note="assembly_fragment:01817
fragment_chain:3"
41202..44151
/note="assembly_fragment:00825
fragment_chain:4"
44252..48108
/note="assembly_fragment:01678
fragment_chain:4"
48209..58414
/note="assembly_fragment:01296
fragment_chain:5"
58515..77249
/note="assembly_fragment:00448
fragment_chain:5"
77350..82656
/note="assembly_fragment:01760
fragment_chain:6"
82757..85452
/note="assembly_fragment:00800
fragment_chain:6"
85553..90378
/note="assembly_fragment:00196"
90479..95473
/note="assembly_fragment:00312"
95574..99853
/note="assembly_fragment:00433"
98654..105535
/note="assembly_fragment:00502"
```

```
misc_feature 105636..114741
/note="assembly_fragment:00551"
misc_feature 114842..118544
/note="assembly_fragment:00557"
misc_feature 118645..130337
/note="assembly_fragment:00701"
misc_feature 130438..144196
/note="assembly_fragment:00873"
misc_feature 144297..149313
/note="assembly_fragment:01103"
misc_feature 149414..152428
/note="assembly_fragment:01347"
misc_feature 152529..154910
/note="assembly_fragment:01545"
misc_feature 155011..163649
/note="assembly_fragment:01263
fragment_chain:7"
misc_feature 163750..166876
/note="assembly_fragment:01448
fragment_chain:7
clone_end:SP6
vector_side:right"

BASE COUNT 51122 a 33170 c 31820 g 48259 t 2505 others

ORIGIN

Alignment Scores:
Pred. No.: 8.99e+03 Length: 166876
Score: 46.00 Matches: 9
Percent Similarity: 76.47% Conservative: 4
Best Local Similarity: 52.94% Mismatches: 4
Query Match: 56.10% Indels: 0
DB: 2 Gaps: 0

US-10-081-935-2 (1-18) x AL356864 (1-166876)

Oy 1 MetClnPheserHIsAlaleuJlAlaleuValAlaGlyLeuAlaSer 17
Db 68948 CTTCAACACATCATATGCCCTAGTACTTGTGTACGAGAAACCTTGCACT 68898

RESULT 37
AC118440/c 169693 bp DNA linear HTG_19-NOV-2002
LOCUS Rattus norvegicus clone CH230-255F21, WORKING DRAFT SEQUENCE, 6
DEFINITION
AC118440
AC118440.4 GI:25099817
VERSION
AC118440
AC118440
KEYWORDS
HTG: HTGS PHASE1: HTGS DRAFT: HTGS_FULLTOP.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 169693)
REFERENCE
Muzny D,Marie, Metzker M,Lee, Abramson S., Adams C., Alder J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Angiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Bewalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Caldwell,E.,
Cardenas,V., Carter,K., Cavazos,I., Cessari,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Devila,M.L., Davis,C., Davy-Carrillo,L., De Anda,C., Dedrich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabiri,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregiorgis,E., Geer,K., Gili,R., Grady,M., Guerra,M., Guevara,W.,
Gunaratne,P., Haaland,M., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A.,
```


Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Koyar, C., Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, J., Liu, J., Liu, D., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuwa, L., Louised, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., McInnes, S., McLeod, M., McNeill, T., Meenen, E., Mitoavljic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokileh, O., Okunolu, G., Olarunsgoon, A., Pal, S., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C., Smajs, D., Sneed, A., Sodergren, E., Song, X., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Swatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zinn, D., von Niederhausen, A., Weiss, R., Smith, D., Holt, R., Smith, H., O., Weinstock, G., and Gibbs, R. A.

Unpublished
2 (bases 1 to 169693)
Direct Submission
Submitted (17-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 169693)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:23194840.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/atlas/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GUVZ

Center clone name: CH230-255F21

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 161764 bases at least Q40

Consensus quality: 163810 bases at least Q30

Consensus quality: 165085 bases at least Q20

Estimated insert size: 165089; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

• (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
• NOTE: This is a 'working draft' sequence. It currently
• consists of 6 contigs. The true order of the pieces
• is not known and their order in this sequence record is
• arbitrary. Gaps between the contigs are represented as
• runs of N, but the exact sizes of the gaps are unknown.
• This record will be updated with the finished sequence
• as soon as it is available and the accession number will
• be preserved.

1 71663: contig of 71663 bp in length
* 71664 71763: gap of unknown length
* 71764 162672: contig of 90909 bp in length
* 162673 162772: gap of unknown length
* 162773 164075: contig of 1303 bp in length
* 164076 164175: gap of unknown length
* 164176 165368: contig of 1193 bp in length
* 165369 165469: gap of unknown length
* 165469 167978: contig of 2510 bp in length
* 167979 168078: gap of unknown length
* 168079 169693: contig of 1615 bp in length.
Location/Qualifiers
1. 169693
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-255F21"
66802..66880
/note="wgs contig"
69386..71663
/note="wgs contig"
complement(788241..88650)
/note="clone_boundary
clone_end:17
site:
end sequence: B2115992"

BASE COUNT 39475 a 42528 c 43376 g 40890 t 3424 others
ORIGIN

Alignment Scores:
Pred. No.: 9.14e+03 Length: 169693
Score: 46.00 Matches: 10
Percent Similarity: 86.67% Conservative: 3
Best local Similarity: 66.67% Mismatches: 2
Query Match: 56.10% Indels: 0
DB: 2 Gaps: 0

US-10-081-935-2 (1-18) x AC118440 (1-169693)

QY 4 SerH1e1A1e1u1e1A1e1u1A1A1e1G1y1e1u1A1e1S1e1r1A1a 18
Db 90336 TCACACCCCTCATCACGCTTGCTGCGGTGTGAGGTGCA 90292
|||||

RESULT 38
AP001548/c 170270 bp DNA linear HTG 30-MAY-2000
LOCUS

DEFINITION Homo sapiens chromosome 18 clone RP11-83712 map 18q21, WORKING
AP001548
DRAFT SEQUENCE, 17 unordered pieces.
AP001548
AP001548.2 GI:8117386
VERSION
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 170270)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Tokoki, Y., Watanabe, H. and Sakaki, Y.
Homo sapiens 170,270 genomic DNA of 18q21
2 Published Only in Database (2000)
2 (bases 1 to 170270)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Tokoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission

JOURNAL

Submitted (24-MAR-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923, Fax:81-42-778-9924)

COMMENT

On May 30, 2000 this sequence version replaced gi:7340850.

----- Genome Center

Center: RIKEN Genomic Sciences Center (GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

----- Project Information

Center project name: HumDrafl18

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 162012 bases at least Q40

Consensus quality: 165830 bases at least Q30

Insert size: 168670; sum-of-contigs

Quality coverage: 5.46x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1      30841 contig of 30841 bp in length
30942 49995 contig of 19054 bp in length
50096 70667 contig of 20572 bp in length
70768 89290 contig of 18523 bp in length
89391 105447 contig of 16057 bp in length
105548 120373 contig of 14826 bp in length
120474 133152 contig of 12679 bp in length
133253 143203 contig of 9951 bp in length
143304 147902 contig of 4599 bp in length
148003 152506 contig of 4504 bp in length
152607 155699 contig of 3093 bp in length
156806 158705 contig of 2906 bp in length
158806 162039 contig of 3234 bp in length
162140 164114 contig of 1975 bp in length
164215 166502 contig of 2288 bp in length
166603 169081 contig of 2479 bp in length
169182 170270 contig of 1069 bp in length

```

Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1      30841: contig of 30841 bp in length
30842 30941: gap of 100 bp
30942 49995: contig of 19054 bp in length
49996 50095: gap of 100 bp
50096 70667: contig of 20572 bp in length
70668 70768: gap of 100 bp
70768 89290: contig of 18523 bp in length
89291 89391: gap of 100 bp
89391 105447: contig of 16057 bp in length
105448 105548: gap of 100 bp
105548 120373: contig of 14826 bp in length
120374 120474: gap of 100 bp
120474 133152: contig of 12679 bp in length
133153 133253: gap of 100 bp
133253 143203: contig of 9951 bp in length

```

```

*      143204 143303: gap of 100 bp
*      143303 147902: contig of 4599 bp in length
*      147902 148003: gap of 100 bp
*      148003 152506: contig of 4504 bp in length
*      152507 155699: gap of 100 bp
*      155699 158705: contig of 3093 bp in length
*      158705 158806: gap of 100 bp
*      158806 162039: contig of 3234 bp in length
*      162039 164114: gap of 100 bp
*      164114 164215: contig of 1975 bp in length
*      164215 166502: contig of 2288 bp in length
*      166502 169081: contig of 2479 bp in length
*      169081 169182: gap of 100 bp
*      169182 170270: contig of 1069 bp in length.

```

FEATURES

source

```

1..170270
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18q21"
/clone="RP11-83712"

```

misc_feature

```

1..30841
/note="assembly_fragment"

```

misc_feature

```

30942..49995
/note="assembly_fragment"

```

misc_feature

```

50096..70667
/note="assembly_fragment"

```

misc_feature

```

70768..89290
/note="assembly_fragment"

```

misc_feature

```

89391..105447
/note="assembly_fragment"

```

misc_feature

```

105448..120373
/note="assembly_fragment"

```

misc_feature

```

120474..133152
/note="assembly_fragment"

```

misc_feature

```

133253..143203
/note="assembly_fragment"

```

misc_feature

```

143304..147902
/note="assembly_fragment"

```

misc_feature

```

148003..152506
/note="assembly_fragment"

```

misc_feature

```

152607..155699
/note="assembly_fragment"

```

misc_feature

```

155806..158705
/note="assembly_fragment"

```

misc_feature

```

158806..162039
/note="assembly_fragment"

```

misc_feature

```

162140..164114
/note="assembly_fragment"

```

misc_feature

```

164215..166502
/note="assembly_fragment"

```

misc_feature

```

166603..169081
/note="assembly_fragment"

```

misc_feature

```

169182..170270
/note="assembly_fragment"

```

BASE COUNT

```

51510 a 32818 c 32593 g 51749 t 1600 others

```

ALIGNMENT

Scores:

```

Pred. No.: 9.17e+03
Score: 46.00
Percent Similarity: 93.75%
Best Local Similarity: 43.75%
Query Match: 56.10%
DB: 2
Gaps: 0

```

```

US-10-081-935-2 (1-18) x AP001548 (1-170270)

```

QY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaIleAglyLeuAla 16
LOCUS AC090405 170413 bp DNA linear HTG 27-MAY-2001
DEFINITION Homo sapiens chromosome 18 clone RP11-779H19 map 18, WORKING DRAFT
AC090405
ACCESSION AC090405.2 GI:14210571
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 170413)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barnett,N., Bostien,V., Boguslavsky,L., Boukhgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodgson,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardner,S., Ginde,S., Goyette,M., Graham,L., Grand-pierre,N.,
Hagood,B., Heathford,A., Horton,L., Hulme,W., Hulten,I., Johnson,R.,
Jones,C., Karasik,A., Labrecque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., Maclean,C., MacDonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Riebeck,W., Riley,R., Rise,C., Rogov,P., Roman,J., Roselli,M.,
Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testave,S., Theodore,J.,
Travers,M., Travis,N., Triggillo,J., Vasilev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zemdek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 27, 2001 this sequence version replaced gi:12958041.
All repeats were identified using RepeatMasker:
Smit,A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12714
Center clone name: 779.H.19
----- Summary Statistics
Sequencing vector: Plasmid, n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 165803 bases at least Q40
Consensus quality: 16767 bases at least Q30
Consensus quality: 168509 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 169313; sum-of-contents
Quality coverage: 8.9 in Q20 bases; agarose-fp
Quality coverage: 8.3 in Q20 bases; sum-of-contents

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 27912: contig of 27912 bp in length
* 27913 28012: gap of 100 bp
* 28013 28759: contig of 747 bp in length
* 28760 28859: gap of 100 bp
* 28860 29596: contig of 737 bp in length
* 29597 29696: gap of 100 bp
* 29697 30720: contig of 1024 bp in length
* 30721 30820: gap of 100 bp
* 30821 31522: contig of 702 bp in length
* 31523 31622: gap of 100 bp
* 31623 32368: contig of 746 bp in length
* 32369 32468: gap of 100 bp
* 32469 33330: contig of 862 bp in length
* 33331 33430: gap of 100 bp
* 33431 34637: contig of 1207 bp in length
* 34638 34737: gap of 100 bp
* 34738 38948: contig of 4211 bp in length
* 38949 39048: gap of 100 bp
* 39049 43570: contig of 4522 bp in length
* 43571 43670: gap of 100 bp
* 43671 162170: contig of 118500 bp in length
* 162171 162270: gap of 100 bp
* 162271 170413: contig of 8143 bp in length.
FEATURES
source
1. 170413
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-779H19"
/clone_1lb="RPC1-11 Human Male BAC"
1. 27912
/note="assembly_fragment"
clone_end:Spe
vector_side:left
28013..28759
/note="assembly_fragment"
28860..29596
/note="assembly_fragment"
29697..30720
/note="assembly_fragment"
30821..31522
/note="assembly_fragment"
31623..32368
/note="assembly_fragment"
32469..33330
/note="assembly_fragment"
33431..34637
/note="assembly_fragment"
34738..38948
/note="assembly_fragment"
39049..43570
/note="assembly_fragment"
43671..162170
/note="assembly_fragment"
162271..170413
/note="assembly_fragment"
vector_end:T7
BASE COUNT 50853 a 32105 c 32644 g 53709 t 1102 others
ORIGIN

Alignment Scores:
Pred. No.: 9,18e+03 Length: 170413
Score: 46.00 Matches: 7
Percent Similarity: 93.75% Conservative: 8
Best Local Similarity: 43.75% Mismatches: 1
Query Match: 56.10% Indels: 0

DB: 2 Gaps: 0

US-10-081-935-2 (1-18) x ACC090405 (1-170413)

Qy 1 MetGlnPheSerHisAlaLeuIleAlaValAlaIaGlyLeuAla 16

Db 131683 TTGAGTTTACTACTCTCTCTCTCTCTCTCTCTCTGACGCTGGAATGCCA 131730

RESULT 40

AC134519 174171 bp DNA linear HTG 27-SEP-2002

LOCUS Homo sapiens chromosome 5 clone RP11-659C12, WORKING DRAFT

DEFINITION SEQUENCE:

AC134519 AC134519.1 GI:23334701

VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS 1 (bases 1 to 174171)

TITLE DOE Joint Genome Institute.

JOURNAL Sequencing of Human Chromosome 5

REFERENCE 2 (bases 1 to 174171)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (27-SEP-2002) Production Sequencing Facility, DOE Joint

COMMENT Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Center: Joint Genome Institute

Center Code: JGI

Web site: http://www.jgi.doe.gov

Project Information

Center Project Name: 1547620

Center clone name: RPCT-11_659C12

Summary Statistics

Consensus quality: 173858 bases at least Q40

Consensus quality: 174094 bases at least Q30

Consensus quality: 174162 bases at least Q20

Estimated insert size: 170000; agarose-fp estimation

Estimated insert size: 174171; sum-of-contigs estimation

Quality coverage: 11.95 in Q20 bases; agarose-fp estimation

Quality coverage: 11.67 in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 174171: contig of 174171 bp in length.

Location/Qualifiers

1..174171

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="5"

/clone="RP11-659C12"

/clone_lib="RPCT human BAC library 11"

BASE COUNT 55289 a 30550 c 30653 g 57679 t

ORIGIN

US-10-081-935-2 (1-18) x AC134519 (1-174171)

Qy 2 GlnPheSerHisAlaLeuIleAlaValAlaIaGlyLeuAlaSer 17

Db 121346 AATACCATCATGCGCTATTAGCAAAAGTGGCACTTGCTTATCTCA 121393

Search completed: November 13, 2003, 12:34:37

Job time : 2171 secs

Alignment Scores:

Pred. No.: 9.39e+03 Length: 174171

Score: 46.00 Matches: 9

Percent Similarity: 81.25% Conservative: 4

Best Local Similarity: 56.25% Mismatches: 3

Query Match: 56.10% Indels: 0

DB: 2 Gaps: 0

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2003, 00:08:38 / Search time 2382 Seconds
(without alignments)
3331.850 Million cell updates/sec

Title: US-10-081-935-4
Sequence: 1023
1 MOPSHALIALVAGLASAQL.....NVRASVGVIAALLGLAAVL 194

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n.model -DEV=x1h
-Q/cg2_1/USFTO.spool/US10081935/runat_04122003_134829_10475/app_query.fasta_1.391
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=pro -NORM=exc -HEAFSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10081935 @CGN 1.1.2372 @runat_04122003_134829_10475 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -KGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sbs.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_ov.*
22: em_or.*
23: em_pal.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sbs.*
28: em_un.*

29: em_vi.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rtd.*
36: em_hcg_mam.*
37: em_hcg_vtr.*
38: em_by.*
39: em_hgo_hum.*
40: em_hgo_mus.*
41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	1023	100.0	1201	8	CIU39835	U39835 Coccidioid
2	1023	100.0	1234	8	CIU32518	U32518 Coccidioid
3	970	94.8	1435	8	CIU51200	US1200 Coccidioid
4	970	94.8	3821	8	AF013256	AF013256 Coccidiol
5	206.5	20.2	97478	2	AC105929	AC105929 Magnapor
6	186	18.2	2960	3	AY057052	AY057052 Mamestra
7	185.5	18.1	1470	6	AR140164	AR140164 Sequence
8	184	18.0	149696	14	AF19782	AF19782 Gallitric
9	182.5	17.8	3655	3	DD1GSPR	M3861 Dictyosteli
10	175	17.1	2455	3	TNAF000605	AF000605 Trichoplu
11	175	17.1	2455	6	AR129835	AR129835 Sequence
12	175	17.1	2821	3	TNAF000606	AF000606 Trichoplu
13	175	17.1	2821	6	AR129836	AR129836 Sequence
14	172.5	16.9	1011	1	CTHA20770	CTHA20770 Clostridi
15	172.5	16.9	3067	1	CTCH1ACGN	Z68934 C. thermocel
16	170.5	16.7	207870	2	AC115017	AC115017 Mus muscu
17	170	16.6	5437	1	CASR69XYN2	AF036924 Caldicell
18	169	16.5	3114	8	AF093132	AF093132 Pneumocys
19	167	16.3	256635	2	AC128838	AC128838 Rattus no
20	166	16.2	1500	8	AY166602	AY166602 Magnapor
21	165.5	16.2	4291	8	AF080221	AF080221 Candida a
22	165.5	16.2	4792	8	AF254147	AF254147 Candida a
23	165.5	16.2	238046	2	AC095003	AC095003 Rattus no
24	164.5	16.1	225849	2	AC098905	AC098905 Rattus no
25	164	16.0	4675	3	DDU20608	U20608 Dictyosteli
26	163.5	16.0	110000	3	AC116957_0	AC116957 Dictyoste
27	162.5	15.9	222896	2	AC119773	AC119773 Rattus no
28	162	15.8	4284	3	DDU20661	U20661 Dictyosteli
29	161.5	15.8	252801	2	AC095211	AC095211 Rattus no
30	161.5	15.8	300029	1	AE012556	AE012556 Xylella f
31	161	15.7	9435	3	AF269242	AF269242 Plasmodiu
32	160.5	15.7	110000	2	AC109525_0	AC109525 Rattus no
33	160.5	15.7	237019	2	AC094443	AC094443 Rattus no
34	160	15.6	124636	9	AP000721	AP000721 Homo sapi
35	159.5	15.6	259772	2	AC106943	AC106943 Rattus no
36	158	15.4	9636	6	PFAP8230X	L08135 Plasmodium
37	158	15.4	9636	6	AR178096	AR178096 Sequence
38	158	15.4	9636	6	195887	L04152 Plasmodium
39	158	15.4	9654	3	PFAS230A	AE001393 Plasmodiu
40	158	15.4	15148	3	AE001393	AE001393 Plasmodiu
41	158	15.4	251908	2	AC094522	AC094522 Rattus no
42	158	15.4	252301	2	AC111805	AC111805 Rattus no
43	157.5	15.4	4034	14	TTV1TP	X14717 Thermoprote
44	157.5	15.4	13669	14	TTV1TP	X14855 Thermoprote
45	157	15.3	1784	8	AF254145	AF254145 Candida a

RESULT 1

ALIGNMENTS

LOCUS	CUJ39835	1201 bp	mRNA	linear	PLN 18-JAN-1996
DEFINITION	Coccidioides immitis immunoreactive cell wall protein mRNA, complete cds.				
ACCESSION	U39835				
VERSION	U39835.1	GI:1161372			
KEYWORDS					
SOURCE	Coccidioides posadasii				
ORGANISM	Coccidioides posadasii Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onygenales; mitosporic Onygenales; Coccidioides.				
REFERENCE	1 (bases 1 to 1201) Dugger,K.O., Villareal,K.M., Nguyen,A., Zimmermann,C.R., Law,J.H. and Galgiani,J.N. Cloning and sequence analysis of the cDNA for a protein from Coccidioides immitis with immunogenic potential Biochem. Biophys. Res. Commun. 218 (2), 485-489 (1996)				
TITLE					
JOURNAL					
MEDLINE	96144702				
PUBMED	8561782				
REFERENCE	2 (bases 1 to 1201) Dugger,K.O. and Galgiani,J.N. Direct Submission Submitted (01-NOV-1995) Kirs O. Dugger, Infectious Diseases, University of Arizona, 1501 N. Campbell, Tucson, AZ 85724, USA				
AUTHORS					
TITLE					
JOURNAL					
FEATURES	location/Qualifiers				
source	1..1201				

CD:

```

"note="33 kDa apoglycoprotein"
/codon_start=1
/product="immunoreactive spherule cell wall protein"
/protein_id="AAC49140.1"
/db_xref="GI:1161374"
/translation="MQSFHALIALVAAGLASAQIPDIPPCALNCFEALGNDGCTRLT
DFKCHSKPELPQGITPCVEEALCPDAISIVSNIVQCSAGVPIIDIPVDITAAPE
PSEIAEPTEPEPTPEPTAEPAAPEHPHEPTPTVPTGTGGVPTGCSFTVTGR
PTASTPAEPFPAAGSNVRAVSVGIIAALGLAAYL"

```

Alignment Scores:	
Pred. No.:	1 57e-51
Score:	1023.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	8
	Gaps: 0
	Length: 1201
	Matches: 194
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

US-10-081-935-4 (1-194) X CIU39835 (1-1201)

QY	1	MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu	20
Db	152	ATGAGATTCTCTCAGCTCTCATGCTCTCGCTCGCTGGCGGCGCTTCGACAGGCCAGCTTC	211
QY	21	ProAspIleProProCysAlaLeuAsnCysPheValGluAlaLeuGluAsnAspGlyCys	40
Db	212	CCAGGACATCCACGCTTGCGCTTCGAACGCTTCGTTGAGGCTCTTCGGACAGATGGCTGC	271
QY	41	ThrArgLeuThrAspPheLeuGlyCysHisCysSerLysProGluLeuProGlyGlnIleThr	60
Db	272	ACTGCGTTGACCGCACTTCMACTGCCACAGCTCCAAAGCTGAGCTCCCGAGGACAGTCACT	331
QY	61	ProCysValGluGluAlaCysProLeuAspAlaArgIleSerValSerAsnIleValAla	80
Db	332	CCTTGCGTTGAGGAGGCGCTGCCCTTCGACCGCCGCTATCTCCGTTCCAAATGCTGCTT	391
QY	81	AspGlnCysSerLysAlaGlyValProIleAspIleProProValAspThrThrAlaAla	100
Db	392	GACCAGTCTCTCAAGGCGCGGTGTCCCAATTGACATCCACCAAGTTGACACCAACGCGCT	451

Accession	Protein	Length
QY	ProGluProSerGluThrAlaGluProThrAlaGluProThrGluProThrAlaGlu	120
Db	452 CCGAGCCATCGAGACCGCTGAGCCACCGCTGAGCCACCGAGGCCCATCGCAG	511
QY	121 ProThrAlaGluProThrAlaGluProThrAlaGluProThrGluGluProThrAlaVal	140
Db	512 CCTACCGCTGAGCCACCGCTGAGCCACTCATGAGCCACCGAGAGCCCATCGCGTC	571
QY	141 ProThrGlyThrGlyGlyValProThrGlyThrGlySerPheThrAlaThrGlyArg	160
Db	572 CCACCGGCACCTGGCGTGTGTCTCCACTGGCACCAGGTTCTTCACTCCCTACTGCAGA	631
QY	161 ProThrAlaSerThrProAlaGluPheProGlyAlaGlySerAsnValAlaArgAlaSerVal	180
Db	632 CCAACTGGCTCCACCCACCGACTGAGTTCACAGGTGTGTCTCCAAAGTCGTCGCACGCGTT	691
QY	181 GlyGlyIleAlaAlaAlaLeuLeuGlyIleAlaAlaIleTyrLeu	194
Db	692 GCGCGCATTCGCTGCTCTCTCTCGGTCTCGCTGCTCACTG 723	

RESULT 2

LOCUS	1234 bp	mRNA	linear	PLN 04-NOV-1996
CTU32518				
Coccidioides immitis				
Coccidioides antigen 2 (Ag2)		mRNA,	complete	
DEFINITION		cds.		

VERSION	U32518.1	GI:1200179
KEYWORDS	.	
SOURCE	Coccidioides posadaei	
ORGANISM	Coccidioides posadaei	

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 1231)	Zhu, Y., Yang, C., Magee, D. M. and Cox, R. A.	Molecular cloning and characterization of <i>Coccidioides immitis</i> antigen 2 cDNA

MEDLIN

PUBMED	8698497
REFERENCE	2 (bases 1 to 1234)
AUTHORS	Cox, R.A.

JOURNAL

Research Immunology, 2303 S.E. Military Drive, San Antonio, TX 78223, USA
On Feb 24, 1996 this sequence version replaced gi:1184058.

son.

```

/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="Silveira"

```

gen

CDS

```

/note="major T-cell-reactive component of mycelia and
spherule cell walls"
/codon_start=1
/product="antigen 2"
/protein_id="AAB18148.1"
/db_xref="GI:1200180"
/translation="MQSHALILVLAAGLSAQLPDIPCALNCPEVLEKNGD
DFKCHPELPQGITVCVEACPLDARIISVNI VVDQSKAVPIDIPVDI
PSEIAEPKAPTEPTAEPAPTEPTHEPTPEPTAIPVPTGGVPTGSGS
PTASTPEFPAGSANSVASVSGIIAALLGLAAYL"
polya_site
1234

```

BASE COUNT	252 a	383 c	277 g	322 t
ORIGIN				

Alignment Scores:

Db	835	CCAACTGCGTCCACCCCAGCTGAAGTTCCACAGTGCTGCATCAAGTCGTGCCAGAGGTT	894
Oy	181	GlyGIYtIleAlaAAlaLeuLcUgLYleuAlaAlaTYrLeu	194
Db	895	GGCGGCATTGCTGCTGCTCTCCTCGCTCGCTGCTGCTACTG	936
RESULT 4			
LOCUS	AF013256	3821 bp	DNA linear PLN 22-AUG-2000
DEFINITION	Coccidioides immitis proline rich antigen (PRA) gene, complete cds.		
ACCESSION	AF013256		
VERSION	AF013256.1	GI:2331288	
KEYWORDS			
SOURCE			
ORGANISM	Coccidioides posadasii Coccidioides posadasii Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onygenales; mitosporic Onygenales; Coccidioides. 1 (bases 1 to 3821) Peng,T., Orsborn,K.I., Orbach,M.J. and Galgiani,J.N. Proline-rich vaccine candidate antigen of Coccidioides immitis: conservation among isolates and differential expression with sporule maturation J. Infect. Dis. 179 (2), 518-521 (1999)		
JOURNAL	PENG,T., ORSBORN,K.I., ORBACH,M.J. AND GALGANI,J.N. Submitted (09-JUL-1997) Infectious Disease, University of Arizona, 1501 N. Campbell, Tucson, AZ 85724, USA Location/Qualifiers 1..3821 /organism="Coccidioides posadasii" /mol_type="genomic DNA" /strain="Silveria" /db_xref="taxon:199306"		
MEDLINE	99094977 9878042 2 (bases 1 to 3821) Peng,T., Orsborn,K.I., Orbach,M.J. and Galgiani,J.N. Direct Submission Submitted (09-JUL-1997) Infectious Disease, University of Arizona, 1501 N. Campbell, Tucson, AZ 85724, USA Location/Qualifiers 1..3821 /organism="Coccidioides posadasii" /mol_type="genomic DNA" /strain="Silveria" /db_xref="taxon:199306"		
PUBMED	99094977 9878042 2 (bases 1 to 3821) Peng,T., Orsborn,K.I., Orbach,M.J. and Galgiani,J.N. Direct Submission Submitted (09-JUL-1997) Infectious Disease, University of Arizona, 1501 N. Campbell, Tucson, AZ 85724, USA Location/Qualifiers 1..3821 /organism="Coccidioides posadasii" /mol_type="genomic DNA" /strain="Silveria" /db_xref="taxon:199306"		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
gene			
mRNA			
CDS			
BASE COUNT	924 A 981 C 903 G 1013 T		
ORIGIN			
Alignment Scores:			
Pred. No.:	5.68e-48	Length:	3821
Score:	970.00	Matches:	193
Percent Similarity:	76.38%	Conservative:	1
Best Local Similarity:	75.98%	Mematches:	0
Query Match:	94.82%	Indels:	60
DB:	8	Gaps:	2
US-10-081-935-4 (1-194) x AF013256 (1-3821)			
Oy	1	MetGlnPhesGrHIsalaleuIlleaAlaLeuValaAlaagIyleuAlaSerAlaGlnLeu	20
Db	896	ATGAGATTCTCTCACGCTCATCGCTCTCGCTGCGGCTCGCAGTCCAGGCTC	955
Oy	21	ProbapllePrOPrOCyAla-----	27
Db	956	CCAGACATCCACCCTTGCGCTGATATTTCGAACCTTTTGTGTGAAACCTCAGAGAGA	1015

Oy		28	-----LeuAnCySPheValGluAa	34
Db		1016	ACGAAGAAGACTCTTGAAATGCTAACGGCTTTGATGCTAGCTCAACTGCTTGTGGAGGCT	1075
Oy		35	LeuGlYaaNaApbJlCYarThArglLeuThrAspPheLYaCSyhiSCysSerLYsProGU	54
Db		1076	CTGGCAACCATGAGTCCTCAGCTCGCTTAGCCAGCTTCAAGTGCACCCTCTCCAAACCTTAG	1135
Oy		55	LeuProGLylglnIlethrProCYaSVaIGUcUGuaLacySproLeuAapAlaArgllle---	73
Db		1136	CTCCACAGAACAGATCACTCTTGCTTGAGAGAGGCCCTCTTCGACGCCCGTATCTGT	1195
Oy		73	-----	73
Db		1196	AAGTAAGCCGACACCAATTCTTGACATCTGCGCCACAAGATGGGATAAGATGAGAAGATCG	1255
Oy		74	-----SerValSerAenIIleValYal	80
Db		1256	AAAGAACTCGAAMATGACGAGCTAACATGTTAAAAATCAGCCGCTCTCCACATCGTGGTT	1315
Oy		81	AspGlnCYaserylVsaIagLYvaIProIGlaAspIlleProProVALaspThrThRALaLa	100
Db		1316	GACCAAGTGCCTCAAAGCCGGGTGTCCCAATTGACATCCACACAGTTGACACCAACCCGCGCT	1375
Oy		101	ProGUlProSerGIuThrAlaGUlProThrAlaGUlProThrGIuGUlProThrAlaGUl	120
Db		1376	CCCAGGCATCCGAGACCGCTGAGCCACCGCTGAGCCAACCGAGAGCCACGATCGCGAG	1435
Oy		121	ProThrAlaGUlProThrAlaGUlProThrHiAgIUProThrGIuGUlProThrAlaYal	140
Db		1436	CTTACCCCTGAGCCACCGCTGAGCCGACCTCATGAGCCACCGAGAGCCACGATCGCGTC	1495
Oy		141	ProThrGIyThrGIyGIyGlyValIProThrGIYThrGIYSerPheThrValIThrGIYArg	160
Db		1496	CCAACCGGACATGGCGGGTGGTGTCCCACTGGCACCGGTTCTTCACTGATCAGCGAGA	1555
Oy		161	ProThrAlaserThrProAlaGUlPherProGIYAlaGLYSerAanValArgAlaSerVal	180
Db		1556	CCAACCTGCTCCACCCAGCTGAATCTCCAGTGCTGCTGCTCCAACTGCCGTCGACGCTT	1615
Oy		181	GIYGIYIleAlaIleAlaLeuLeuGLYleuAlaIatYrlleu	194
Db		1616	GGCGGCAITTGCTGCTGCTCTCTCTGATCTGCTGCTGCTTAAGCTG	1657
RESULT 5				
LOCUS	AC105923/c			
DEFINITION	97478 bp DNA linear HTG 11-JAN-2002			
PROGPRSS	***, 7 unordered pieces.			
ACCESSION	AC105929			
VERSION	AC105929.1 GI:18129402			
KEYWORDS	HTG; HTGS PHASE1.			
SOURCE	Magnaporthe grisea			
ORGANISM	Magnaporthe grisea Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes Incertae sedis; Magnaporthaceae; Magnaporthe. 1 (bases 1 to 97478) Dean,R.A. Dr. Mitchell,T. Dr. Thon,M. Dr. Brown,D.E., Taro,A. and Papalas,J.			
REFERENCE				
AUTHORS				
TITLE	Magaporthe Grisea chromosome 7			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 97478)			
AUTHORS	Dean,R.A. Dr. Mitchell,T. Dr. Thon,M. Dr and Brown,D.E.			
TITLE	Direct Submmission			
JOURNAL	Submitted (11-JAN-2002) Plant Pathology - Fungal Genomics Laboratory, North Carolina State University, 840 Main Campus Drive, Raleigh, NC 27606, USA			
COMMENT	* NOTE: This is a 'working draft' sequence. It currently * consists of 7 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence			


```
Db      216 AAGTTT-----TACTACTGTGAATATGCTGTAAGTATATCGAACCCAGAACTGTCT 269
Oy      51 SerLyseProgluubProgluynilethProCyeva1-----gluualaCy5 67
Db      270 TCTGCTACTGACTTCACAGCTGAGTTCAGCTTGTGTTTCATCCAGCTTCAGAGATGC 329
Oy      68 ProleuAsp-----AlaAtg11eServalseraSn1leVala1AspGlnCyserLy5 85
Db      330 TCTTACCGGATCTTCACATTATGACGACGAGTAACTGCTCTCACCACCGACGCTCTT 389
Oy      86 AlaGlyValPro1leAsp1leProProvalAspThrThra1a1aProgluProserGlu 105
Db      390 ACAGCCGACCCACTGACGCTCTTACAGCCGCTCCACTCAGAGCCCTTACCGCTGACCC 449
Oy      106 Thra1aGluProThra1aGluProThra1aGluProThra1aGluProThra1aGluPro 125
Db      450 ACTGCGGCGCCCTACCGCTGACCCAGCCGCGCCCTTACCGCTGACCCAGCCGCTGACCC 509
Oy      126 Thra1aGluProThra1aGluProThra1aGluProThra1aGluProThra1aGluPro 145
Db      510 ACTGCTGACCCACTGACGCGCCCTTACCGCTCTTACGAGCCGCTGACCCAGCCGCTGACCC 569
Oy      146 GlyGlyValProThra1aGluProThra1aGluProThra1aGluProThra1aGluPro 165
Db      570 ACAGCAGCTCTTACTGACGCTCTTACAGCCGCTGCTCCACCGACGCTCTTACTGACGCCCA 629
Oy      166 Pro 166
Db      630 CCA 632

RESULT 7
ARI40164      1470 bp      DNA      linear      PAT 16-JUN-2001
LOCUS      ARI40164
DEFINITION      Sequence 1 from patent US 6207436.
ACCESSION      ARI40164
VERSION      ARI40164.1 GI:14482660
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 1470)
AUTHORS      Bj.o. Blashed,rmvad,M.Bekelund., Hatakeyama,M., Schulein,M. and
TITLES      Nielsen,J.Bech.
JOURNAL      Endo-8-1,4-glucanases from saccharochrix
FEATURES      Patent: US 6207436-A 1 27-MAR-2001;
SOURCE      Location/Qualifiers
1..1470
/organism="unknown"
BASE COUNT      242 a 563 c 489 g 176 t
ORIGIN

Alignment Scores:
Pred. No.:      0.0143      Length:      1470
Score:      185.50      Matches:      44
Percent Similarity:      52.17%      Conservative:      42
Best Local Similarity:      47.83%      Mismatches:      4
Query Match:      18.13%      Indels:      2
DB:      Gaps:      1

US-10-081-935-4 (1-194) x ARI40164 (1-1470)
Oy      83 CysSerLyseVala---GlyValPro1leAsp1leProProvalAspThrThra1a1aPro 101
Db      424 TGCAACGTCGCGTGGTGGCCGACGACGAGCCGACCCGCGAGCCGCGAGCC 483
Oy      102 GluProserGluThra1aGluProThra1aGluProThra1aGluProThra1aGluPro 121
Db      484 ACGCCCGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGG 543
Oy      122 Thra1aGluProThra1aGluProThra1aGluProThra1aGluProThra1aGluPro 141
Db      544 ACGCCGAGCCGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGG 603
```

```
Oy      142 Thra1aGluProThra1aGluProThra1aGluProThra1aGluProThra1aGluPro 161
Db      604 ACCGCCGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCG 663
Oy      161 oThra1aSerThra1aGluProThra1aGluProThra1aGluProThra1aGluPro 172
Db      664 CAGTGTACCGGCGCTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 697

RESULT 8
AF319782/c
LOCUS      AF319782
DEFINITION      Callictrichine herpesvirus 3 strain Cj0149, complete genome.
ACCESSION      AF319782
VERSION      AF319782.2 GI:24762296
KEYWORDS
SOURCE      Callictrichine herpesvirus 3
ORGANISM      Callictrichine herpesvirus 3
REFERENCE      1 (bases 1 to 105145)
AUTHORS      Cho,Y., Ramer,J., Rivallier,P., Quink,C., Garber,R.L., Beier,D.R.
TITLES      An Epstein-Barr-related herpesvirus from marmoset lymphomas
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 98 (3), 1224-1229 (2001)
MEDLINE      21107697
PUBMED      11158621
REFERENCE      2 (bases 1 to 149696)
AUTHORS      Rivallier,P., Cho,Y.G. and Wang,F.
TITLES      Complete genomic sequence of an Epstein-Barr virus-related
JOURNAL      herpesvirus naturally infecting a new world primate: a defining
MEDLINE      J. Virol. 76 (23), 12055-12068 (2002)
PUBMED      22302315
JOURNAL      3 (bases 1 to 105145)
MEDLINE      12414947
PUBMED      12414947
REFERENCE      Wang,F.
AUTHORS      Direct Submission
TITLES      Submitted (12-APR-2001) Medicine, Brigham & Women's Hospital,
JOURNAL      Harvard Medical School, Channing Laboratory, 181 Longwood Avenue,
1 (bases 1 to 149696)
AUTHORS      Rivallier,P., Cho,Y. and Wang,F.
TITLES      Direct Submission
JOURNAL      Submitted (04-JUN-2002) Medicine, Brigham & Women's Hospital,
1 (bases 1 to 149696)
AUTHORS      Harvard Medical School, Channing Laboratory, 181 Longwood Avenue,
1 (bases 1 to 149696)
TITLES      Sequence update by submitter
JOURNAL      On Nov 7, 2002 this sequence version replaced gi:1367640.
1 (bases 1 to 149696)
TITLES      Location/Qualifiers
1..149696
/organism="Callictrichine herpesvirus 3"
/mol_type="genomic DNA"
/strain="Cj0149"
/species="herpesvirus 3"
/db_xref="taxon:106331"
/ncbi="isolated from lymphoma"
/complement(join(138637..138818,138904..139005,
139106..139219,139294..139458,139572..139661,
139744..139992,140075..140167,140247..140485,959..9661)
/product="C7"
/complement(join(138637..138818,138904..139005,
139106..139219,139294..139458,139572..139661,
139744..139992,140075..140167,140247..140485,959..9661)
/codon_start=1
/product="C7"
/protein_id="AA064290.1"
/db_xref="GI:24762310"
/translation="MAGHWYEVYIPGLFLCPLILPSLFWICSLTFLVGHANIVASV
LFLVLANCLILANWNTTRDPVSGRRSSMSLSVAASATATAMPASFITLSFDGGLL
FGTALVITITVIVLVVMEITVIMKRYVHFVITTLFLSPITLISVACILIOSGALL
IEAVVTTITVLAIFLWLPPOGAEADGTLTLTLNLTALCVLIVLITVLPDAQITLVF
CLPCOMTLFICLIRMICNWRGLTRIIICLKFLYGLISLSLFGWVAFLKEVTLPTT
```

gene

mRNA

CDS

ATVDPRLPLFLFLSSVLVLAIMMEFOTSSLPALFVLIAGMLCTVGVIFLLAG
 VKPLSGMIGCASGITMLVLGVLLVLCSTRASTRESIYEDLRYPTRDANGEXVNGVPP
 RGDADRHRLGEPYDDVEQAT"
 484. .1829
 /gene="C1"
 /join(484. .721,811. .897,1087. .1829)
 /gene="C1"
 /product="C1"
 /note="positional and functional similarity to EBV LMP1;
 transforming gene"

CDS

/codon_start=1
 /product="C1"
 /protein_id="AAK38207.1"
 /db_xref="GI:13676641"
 /translation="MAARRLSPPMLVTULLISTSVALLLFIIFSAGATISTEA
 SLVLLLVTLPLSSNGLOPALNITQCFLLAADLALVLPITSEDLI1A
 ILVIVLVGTTTTLVGAIGIRARSLFCIFPLSLFTLIALLLGFSMLLVAI
 LFMVLVILILLLLVPIPHPLPTSLFRMKQVSDPTGSDRSPOSHNSLPD
 EEDPKDTKQPLCNMTQGGPPVNGOLLGQAOCPPYPCCHIQHPDGSDSDGKSW
 GAAGEDNGNDNTNNGCGDDYKSMRKEEEDNGNDNTNRRLEDGCGDDK
 SRRPEEDNRKDRIGTKP"
 3943. .4464
 /note="similar to EBV BALF1; similar to bcl-2"
 /codon_start=1
 /product="ORF1"
 /protein_id="AAK38208.1"
 /db_xref="GI:13676642"
 /translation="MTDDVFDNAPNLGAEDENDRVLLTRVMAAYLDDPKGLGT
 EERFLRLIKRLMKKEKRPADIVNSGSAPTTLHGIIKRLMFTTRAIYEDHMDNFRV
 RALTSJAVAVYARRNIPGSDENAGLLLVGAFFLCYLRAMLRLGIGIRARAPFLT
 MRRRVGSECYQO"
 4564. .7947
 /note="similar to EBV BALF2; major DNA binding protein"
 /codon_start=1
 /product="ORF2"
 /protein_id="AAK38209.1"
 /db_xref="GI:13676643"
 /translation="MQSPAPEDNLGSHCKVPCGYLYFPPLDGPPIREASLIGTGYH
 GKALTLPLCLGVNENFSPVNAKRRPDPNTGLRATCTPRDAYVHNATATIPV
 RSGRGLIELCDTREYVGDASPPVDAKETVKEEYFSELDSPATVGAIVTEAKE
 RLYSGNLVILSKREVSIQYASICIPLYDREVPFGVAKSAVFNNKALSKMCEYTL
 YGLAQSLRLRDEPFIAMERQSLHDQIKVPKVALKEFNATLSPSSAMV1DT
 ASCELATSYGPAMIDVQEPPTGIRYDAMPVEGCENEDARLKLHFPNALLAHAA
 OJFSTNSVLYVSGVSKTQGGRTNYSFETIGSLQEGTMDNOROVFSGMGAPD
 VNGSGGPNVSVELWFAFSPNNILRAYVYIQPGQKSSINOSHETGSLVAGAT
 SSVCSICESTPAVCNLTIFRLRDPFPVYVTHRDPIYITGASGAVETDFLKLND
 NFVDRDDQAVDEPRYVWQNLNLERLSLIGDGNCEDEPKTRADPMKLKD
 VQAVDAEVQVQFEMGSAKNVQKDLIKSCYHILQYCNVQAQCTVFLQYFRESM
 TLIDIALPICMCYENDNPGLQSAQEMLKGHYOTLCTNERSLAVDKGLTLAKEVVV
 HEPTSDVADLAAIKGRVWORLTIVRMKVLCPKNIKIKRIVPTGENSELONF
 IKSTLRDNYIINGPYIKLNNYHKVLPETKLSCLYHNFARKKTLDTIPROGDAAE
 YSDLAIFDNGSRRAHEQANV1DIVPDDLITVAKORNNNSILKACQTOPFISILHALV
 PIQIVPAPDYPHVLGGDVKLTSSEDTYQISNMTAKTCTVDTFVSSEARAPIT
 LPLTVNKTGVVNGNSOIFQAGNLGYFMGRGVDRNLPGQSLKKOMGCVSRKKRVE
 APPTGLTISKRPQMSVSFEIKRISTLEPAISOEKDESDVDVVCVKLEVCSCAS
 LTLDAEYVILGRYSILASCVLSLEKTLAAAGVEMTEVARDFDLDSLASHLAAGETI
 SIEEPVILPASGGLSIGSHGISTYVRKKRLGLTGLDY"

CDS

CDS

CDS

/note="similar to EBV BALF3; transport protein"
 /codon_start=1
 /product="ORF3"
 /protein_id="AAK38210.1"
 /db_xref="GI:13676644"
 /translation="WTGMLAAVYSOIYAVSFLGICRHLDDSLDTGAVTNIAIRK
 LRTIIGKQIDKONEVCTALSLSELSHLSNLSILEENKRDNDLEPTLIPASFTAK
 FYGOYSQDVSMYLLINDIEIFIRLNSVFCLSYNNGLNNAEKLGLGFRVSEVPD
 RLYLIDVPCMLHEQAVLPQNGRPKISGSRGLCNHLAVVAPIDEPVGLFENSVRO
 ASLTLKKEQDTRIRIHSREKNTTEAOSLDEKCVFSVAPPEISSEILYNNSSR
 NSISLTGEEETWRSLSKLISRDEEMLEWRCDEYLAAGASVNPPIFVSRMPIILRL
 CGGIFSSVGDSEALKIDQASAFPSQOPVRIIADORENYITLNNHMSLRKTEVQR
 ENTERPHAYVANOIGNKDSSVIMSDAKKREYLRQLAEGILTKQACIARQGEIMSE

CDS

CDS

CDS

TLARVMGDTVEEMAAWYHIFARAFVRAVEDKRDGTAEIIFENSKYIKTLHFNOT
 LSSEHNALMTMYSLTLGTLGADGMPPEPVNAVLSYCEAGLIIPOHKLPMASLW
 PGIOPAWEIETSFNSFPQDPDLATLQOILABYIEATITVLSYNTKQOSYHLRV
 DEAGLEGANPPDGIYITTESENPLILIKTKGHIFLDIALTLMLQMDGRV"
 9973. .12585
 /note="similar to EBV BALF4; glycoprotein B"
 /codon_start=1
 /product="ORF4"
 /protein_id="AAK38211.1"
 /db_xref="GI:13676645"
 /translation="MAVFIMIFIFMSYVAGARESEPPVSOQNPKEVTVPGSTHQ
 PYEPPFVCEMNNHGDIFRPATISCPNPLQENHGLIIVPKDNIIPYTPVRAAT
 KVYTVTVVNGVADAVTNAHEERSVYDPELNMOTLVQCYSVNRKAVGVYVVD
 RDDKNTLVLRPASGISTGVRRYNSQDPLDYDDGSGMWTYRRRTYNCMTDQAKNS
 SPNPFVTTLQGVEMSPFYNSGNETPNEPATFEIREYKIVTENRQITPKEER
 AFLNKOTYTLMDLARRSSYCLTLMKDDYSVALIRTOGSSFFHTPGTASPVNTTH
 EKLDPQKIEQOVNKTITOEPAKLSNKYDRAEDIOFFRTAGLLAMLPVPMTH
 ITRIDALNTEAQDQKAADOPDPTANSPTPOAPLRNRNREGSDNTNIDVST
 TEGTGKSDILGTDINPATQIOFAYDSLRSGOINRMGLDASAMCOEORONVNLN
 TKINPTTWGTVGVYQVPAKRIQDVISVRCVPEDESVLSRSMYLSGSETCTYR
 LVSFSFANSTKTEGLOQPNELLTKKAVENCODTQOHYFQSGNEHVRDIOHKT
 IPEVDATLQTFITLNTSFTINDIOTLELYSDERASVDFIEGIFREYNQTONI
 ANLRDLDNAISNNRQPVYDGLDELTDSDGAVQOATNTIVSVGLFSLISGFISSF
 KNIPFGVLLIAGIVAVIPLVIFLRQNRITIAQOVQVFLYPEIOSLTNRREIGOSNAP
 ISKQOLDALMLALEOTAKRTQESKTSSTISLPSRALEAANNRLRLRKKPGRENS
 DLKPLITNEDTEP"
 12590. .15640
 /note="similar to EBV BALF5; DNA polymerase"
 /codon_start=1
 /product="ORF5"
 /protein_id="AAK38212.1"
 /db_xref="GI:13676646"
 /translation="NSGRSFNPNFLHCKNTCAKPKYELRLIPDCRKTGFGAGVTDI
 RGEPEPLCHNGTKIQVMSHGVDGRGMSKCSANGCAVEATDMEFHVDI1EYVT
 QEKCAVIPSDDKQYVPCGIVIKLGRREDEGSCVGINVGOQGYFSETPADVNI
 EV ALVEILKGTGPDRTPCRMATVEVERSSINGDCTRMVYKTLISHPNSISHATSLE
 DKYCKCIFEANDASRRFVLDNKKFTTGGVYRCGRAPRI RDPRSFPDFEDCCTDLS
 VIGNDSMPVWMAFDIECLGEGEPNATNEDTIIQSCVLSMGEPTSYRRILL
 TGTGCAIDGVEYERPSMDLIYAFQILRDLNVTITGINFANDMPTILDRARI
 YGINPASLKITGGGCEVRRPNDACKMPSRACKTKRISGIVIPDMVSVEKELSLSD
 YLDTVAKRI1DGTKEEDVYKSIPIHFRAGPERKRLKGMVCIODSVLWDLNHLFVTH
 VBISEIAKIANIPCRVLDGQOIRVFSCLLAAQOGEYLLPMPESDGRVGCAGATYI
 OPIAGIYVPLVNDPASLYPSIIQHNLCYSYIYVREGGKRLDRPGEDVCSFSL
 GGTFFHVKKHIGFLASLEMLAKRKIRKILGACNBNPROTITDKOLA1KCCN
 ANYGFGVANHGFPCITIAETVTLQGRYTLREAKRVENSLDLEKLAEAKTIFSC
 NEDGDLRV1GDTLSLIECSGSEQVSALEBPKLFTTSLFNAPISLEAKTIFSC
 LMLITRKRVGVLTGKLTIMKGVLDVKTACQFVQRCRVLDLILGNKVKAAASIL
 SLRPYQTSFTEGLPEGFLDVINVLNKAQYDLNANRVPISLSTFTELSTROI1SVKSTO
 MHLIYOKRI1QNRNEPPOIHRIPIPVFAVAGSGSGGTRKSEWADBPVSYTRINI
 PLAVEHYFEKL1QGTANILQCLPGRNPETALSVLONFTAGYPRQO"
 17204. .18121
 /note="similar to EBV BALF1; glycoprotein gp64"
 /codon_start=1
 /product="ORF6"
 /protein_id="AAK38213.1"
 /db_xref="GI:13676647"
 /translation="WDETQIIGSFTEVLEBPQGNSSVNTGLSFTGVMTTEIIL
 ILLFLFLILRLKOSIDNLAVMFMNDTLMLGLKIQFSSVTGLCMLTQHMWMSLL
 LSGFHSNGAARALMWSAKTIKPVNSRCIINYLITLTVILVLAVSFEAGISIC
 NNGPNLCREGATQAGHAIROAKGFEFFCCCAATIIATCTYTIKIHNTSTRMRIC
 NMPTGIIIGATCMLLAALVLPFRPQNGGFCPCPSLMRYTHAFOSLLILLLLLLYI
 WMRPQIDILKNLLESTLVSFRASNRNS"

Alignment Scores:

Pred. No.:	1.25	Length:	149696
Score:	184.00	Matches:	50
Percent Similarity:	46.81%	Conservative:	16
Best Local Similarity:	35.46%	Mismatches:	71
Query Match:	17.99%	Indels:	4
DB:	14	Gaps:	3

US-10-081-935-4 (1-194) x AF319782 (1-149966)

QY 30 CysPheValGluAlaLeuGluGlyAsnAspGlyCysThrArgLeuThrAspPheLeuScyVh1s 49
|||||
2955 TGTTCGCCCATCTATTCGGAAT---CGTCAATCATCATTTGAGTATCGAGTTTGGCC 2899

QY 50 CysSerLeuProGluLeuProGlyGlnIleThrProCysValGluGluAlaCysProLeu 69
|||||
2898 TCTAGTACAGTGAATAGATAGACAGCATTTTGTAGCATGAACTTAGATTCCT 2839

QY 70 AspaAlaGlgIleSerValSerAsnIleValAspGlnCysSerLeuAlaGluValPro 89
|||||
2838 TGTGCTGTTATTATTAGTA-----GTGATATTGTTGTTCTTCTACCGCGCCGACCCCA 2785

QY 90 IleAspIleProProValAspThrThrAlaAlaProGluProSerGluThrAlaGluPro 109
|||||
2784 ACCTCGACCCCAACCTCGACCCCAACCTCGACCCCAACCTCGACCCCAACCTCGACCCCA 2725

QY 110 ThrAlaGluProThrGluGluProThrAlaGluProThrAlaGluProThrAlaGluPro 129
|||||
2724 ACCTCGACCCCAACCTCGACCCCAACCTCGACCCCAACCTCGACCCCAACCTCGACCCCA 2665

QY 130 ThrHisGluProThrGluGluProThrAlaValProThrGlyThrGlyGlyValPro 149
|||||
2664 ACCTCGACCCCAACCTCGACCCCAACCTCGACCCCAACCTCGACCCCAACCTCGACCCCA 2605

QY 150 ThrGlyThrGlySerPheThrValThrGlyArgProThrAlaSerThrProAlaGluPhe 169
|||||
2604 ACCTCGACCCCAACCTCGACCCCAACCTCGACCCCAACCTCGACCCCAACCTCGACCCCA 2548

QY 170 Pro 170
|||
Db 2547 CCA 2545

RESULT 9
DDISGSPB 3655 bp DNA linear INV 02-NOV-1994
LOCUS Dictyostelium discoideum spore germination-specific protein gene,
complete cds.
M33861 J02916
ACCESSION M33861.1 GI:167882
VERSION endo-1,4-beta-D-glucanase; spore germination-specific protein.
KEYWORDS Dictyostelium discoideum
SOURCE Dictyostelium discoideum
ORGANISM Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 3655)
Giorda, R., Ohmachi, T., Shaw, D. R. and Ennis, H. L.
A shared internal threonine-glutamic acid-threonine-proline repeat
defines a family of Dictyostelium discoideum spore germination
specific proteins
JOURNAL Biochemistry 29 (31), 7264-7269 (1990)
MEDLINE 91002566
PUBMED 1976383
REFERENCE 2 (sites)
Blume, J. E. and Ennis, H. L.
A Dictyostelium discoideum cellulase is a member of a spore
germination-specific gene family
JOURNAL J. Biol. Chem. 266 (23), 15432-15437 (1991)
MEDLINE 91332071
PUBMED 1869562
COMMENT Draft entry and computer-readable sequence for [Biochemistry (1990)
in press] kindly submitted
by H.L. Ennis, 17-APR-1990.
FEATURES
source
1..3655
/organism="Dictyostelium discoideum"
/mol_type="Genomic DNA"
/strain="AX-3"
/db_xref="taxon:44689"
join(1346..1412,1505..3555)
/note="spore germination-specific protein"
/codon_start=1

/protein_id="AAA52077.1"
/db_xref="GI:167883"
/translation="MKILKNKCIULLIFGLSLSTOLINADPTDYSLEENALMFYKQNBAG
RUPNDIIPWRGNSALNDASPNKAKDNGDNGSLSGYFDAGDGVKFEFLPMAYSTMTLGM
SFIEVSNIAOGLLSLYLDITIKYGDWLIAMHTANERFAGDVGDDNDVDSWMCPRPD
MTMARPTVMTLTPTEPTEPTEPTEPTEPTEPTEPTEPTEPTEPTEPTEPTEPTEPTEP
YRGVYSDSLITNNOAFPFNSWNGYKDLVMSISIMLYKATQSDSLTKAVADVAGCGVGM
ACGNSHMDMNKXAPCCGLLSKLVPTSTYKTDDEGLNLTWPCGGLNTPYFGGLAMIRQ
WGPARYATLAALFGLSLAGTEKGTDFOKOVDTLIGNPNQDSFVGMGPRYPINPHR
AAHSTTNDINNVNNUYLLKGLVGGPDSNDEYTDRTDISNEVATDYNAGFYCAL
ASLVNSSTSVPT
TERPT
NPDQMSWGHSLTDTNVSSEGTNSISFTPKYAGAVFLGCECIDDTYNNIEFDIN
CGSSGAKOLIRITVYKSKVSKLITDLNGCPTEIANSMTKIKASFIIDPKVSGKVDG
IYIDIKGDTOSTVYISNIATTA"
exon <1346..1412
/number=1
intron 1413..1504
exon 1505..3555
BASE COUNT 1346 a 482 c 459 g 1368 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0495 Length: 3655
Score: 182.50 Matches: 44
Percent Similarity: 45.45% Conservative: 6
Best Local Similarity: 40.00% Mismatches: 49
Query Match: 17.84% Indels: 11
DB: 3 Gaps: 2

US-10-081-935-4 (1-194) x DDISGSPB (1-3655)

QY 84 SerIleValGluValProIleAspIleProProValAspThrThrAlaAlaProGluPro 103
|||||
Db 2806 TCTTCAACTTCTGTTCCACACCACTCCAAGTGAATGAAACCCCAACAGACTCCA 2865

QY 104 SerGlu-----ThrAlaGluProThrAlaGluPro 113
|||
Db 2866 ACTGAGACTCCAACTGAGACTCCAACTGAGACTCCAACTGAGACTCCAACTGAGACTCCA 2925

QY 114 ThrGluGluProThrAlaGluProThrAlaGluProThrAlaGluProThrAlaGluPro 133
|||||
Db 2926 ACAGAGACTCCAACTGAGACTCCAACTGAGACTCCAACTGAGACTCCAACTGAGACTCCA 2985

QY 134 ThrGluGluProThrAlaValProThrGlyThrGlyValProThrGlyThrGly 153
|||||
Db 2986 ACAGAACTCCAACTGAGACTCCAACTGAGACTCCAACTGAGACTCCAACTGAGACTCCA 3045

QY 154 SerPheThrValThrGlyArgProThrAlaSerThrProAlaGluPheProGlyAlaGly 173
|||||
Db 3046 ACTGAACTGTTACTCCAACTGAGACTCCAACTGAGACTCCAACTGAGACTCCA 3102

QY 174 SerAsnValArgAlaSerValGlyIle 183
|||||
Db 3103 GAATCTTATGCAATCTATAAAGTGATTA 3132

RESULT 10
TNAF000605 2455 bp mRNA linear INV 28-JUN-1997
LOCUS Trichoplusia ni
DEFINITION Trichoplusia ni insect intestinal mucin IIM14 mRNA, complete cds.
ACCESSION AF000605
VERSION AF000605.1 GI:2224918
KEYWORDS Trichoplusia ni (cabbage looper)
SOURCE Trichoplusia ni
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Dictyvia;
Noctuidae; Noctuinae; Plusiinae; Trichoplusia.
REFERENCE 1 (bases 1 to 2455)
Wang, P. and Granados, R. R.
Molecular cloning and sequencing of a novel invertebrate intestinal
mucin cDNA

QY 155 PheThValThrglyAtpProThAlaSerThProAlaGluPheProGlyAlaGlySer 174
|||||
Db 1865 ---ACTGTCAATACACCACTACTGTGCCCCCACTACCGCAGCAGCCTGCCCAAGACC 1921

QY 175 AsnVal 176
|||
Db 1922 ACAGTC 1927

RESULT 12
TNAF000606 2821 bp mRNA linear INV 28-JUN-1997
LOCUS TNAF000606
DEFINITION Trichoplusia ni insect intestinal mucin IIM22 mRNA, complete cds.
ACCESSION AF000606
VERSION AF000606.1 GI:2224920
KEYWORDS
SOURCE Trichoplusia ni (cabbage looper)
ORGANISM Trichoplusia ni
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuidae; Noctuidae; Plusiinae; Trichoplusia.
REFERENCE
AUTHORS Wang, P. and Granados, R.R.
TITLE Molecular cloning and sequencing of a novel invertebrate intestinal
mucin cDNA
JOURNAL J. Biol. Chem. 272 (26), 16663-16669 (1997)
MEDLINE 97341213
PUBMED 9195982
REFERENCE 2 (bases 1 to 2821)
AUTHORS Wang, P. and Granados, R.R.
TITLE Direct Submission
JOURNAL Submitted (21-APR-1997) Boyce Thompson Institute, Cornell
University, Ithaca, NY 14853, USA
Location/Qualifiers
FEATURES
source 1..2821
/organism="Trichoplusia ni"
/mol_type="mRNA"
/db_xref="taxon:7111"
26..2449
/codon_start=1
/product="insect intestinal mucin IIM22"
/protein_id="AAC47557.1"
/db_xref="GI:2224921"
/translation="MIKTLFLTLGLVAARPEVSDAKNPALHPHPDPPAEQHWL
LPHEIDCTKRYCEGLKFTIAPRDCAKTEKFSQCTQVHAALAGCTLRGPAETTOA
PATTOAPTTTTOAPTTTOAPTTTOAPTTTOAPTTTOAPTTTOAPTTTOAPTTTOA
PTTTOAPTTTTOAPTTTOAPTTTOAPTTTOAPTTTOAPTTTOAPTTTOAPTTTOA
AATTPAATTPGVPAATPAPVWPICELPNCSPADFDIHLI PHDKYCNLFYQCSNGY
TFEORCPGLYFNRYORCDSPANVECDGEISAPRVEGNEDEDIDGLDNGCPA
NFEIDMLPHGNRCQKYQCVGNLVERCGAGTHFSELOCDHIELVGLTSGES
EELVNDPACTGWCQPTPIEMERLPNGCPADFSIDLPHESOCGOVLOCVHGOTTA
RPGONLHFSRATGSCSPVTAAGCVFECDSNDCTSTAAFTAAFTAAFTAAFT
AASIVPAPATPATAVPPTTAIPPTAIPPTAIPPTAIPPTAIPPTAIPPTAIPPTAIP
AAPTVAPEIDITVSAPTAAPTAAPTAAPTAAPTAAPTAAPTAAPTAAPTAAP
AAPTVAPEIDITVSAPTAAPTAAPTAAPTAAPTAAPTAAPTAAPTAAPTAAP
VAPAPNTAAVTTTSAATPDEDDIDPLPNDIPNCEVCNVLPAHADDCKDYVW
CDGNQYLVCSGEGQNPPTTKTCDFACNVCVRSNIQMSSESYEVQVFIWNKLEDD
IRQALNFEFL"

BASE COUNT 710 a 1030 c 529 g 552 t
ORIGIN

Alignment Scores:
Pred. No.: 0.107 Length: 2821
Score: 175.00 Matches: 56
Percent Similarity: 37.36% Conservative: 12
Best Local Similarity: 30.77% Mismatches: 72
Query Match: 17.11% Indels: 42
DB: 3 Gaps: 7

US-10-081-935-4 (1-194) x TNAF000606 (1-2821)

QY 25 ProCysAlaLeuAenCysPheValGluAlaLeuGlyAsnAspGlyCysThArgLeuThr 44
|||
|||

Db 1406 CTTGCCACACAGTCTGT-----GAGTCTCTGTGACCGCTGTGC-----CAA 1450

QY 45 AspPheLysCys-----HisCysSerLysProGluLeuProGlyGlnLeuThr 60
|||||
Db 1451 GTTTTCGAGTGCAGTTCTGACCAACCACTGACATCTGTGCCGACAGTCTCCA 1510

QY 61 ProCysValGluGluAlaCysProLeuAspAlaArgLieserValSerAsn1LeuVal 80
|||
Db 1511 ACGGCTGCCCCAACGGCTGCCGCCAACACTGCCAACCTCCACCTCCAGCTGTTC 1570

QY 81 AspGlnCysSerLysValGlyValPro1Leasp1LeProProValAspThrThAla 100
|||
Db 1571 CCAGCTGCAACGACCGACCGCAATGACAGCCCACTGCCACT-----ACAACCGCAATT 1624

QY 101 ProGluProSerGluThrAlaGluProThrAlaGluProThrGlu----- 115
|||
Db 1625 CTTACTCCGGCCCCACCGCTGCGCCCAACCGCACTCTACTACTGCTGCTGCAATCC 1684

QY 116 -----GluProThrAlaGluProThrAla----- 123
|||
Db 1685 CCAACCACTGACAGTACCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1744

QY 124 -----GluProThrAlaGluProThrHisGluProThr 134
|||
Db 1745 CCGAATATCCCAATCATCTGATCATAGCGCTACCGCTGCGCCACCGCTGCGCCAC 1804

QY 135 GluGluProThrAlaValProThrGlyThrglyGlyValProThrGlyThrglySer 154
|||||
Db 1805 GCTGCCCCACCGAGCGCCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1852

QY 155 PheThValThrglyAtpProThAlaSerThProAlaGluPheProGlyAlaGlySer 174
|||||
Db 1853 ---ACTGTCAATACACCACTACTGTGCCCCCACTACCGCAGCAGCCTGCCCAAGACC 1909

QY 175 AsnVal 176
|||
Db 1910 ACAGTC 1915

RESULT 13
ARI29836 2821 bp DNA linear PAT 16-MAY-2001
LOCUS ARI29836
DEFINITION Sequence 2 from patent US 6187558.
ACCESSION ARI29836
VERSION ARI29836.1 GI:14117733
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2821)
AUTHORS Granados, R.R. and Wang, P.
TITLE Invertebrate intestinal mucin cDNA and related products and methods
JOURNAL Patent: US 6187558-A 2 13-FEB-2001;
Location/Qualifiers
FEATURES
source 1..2821
/organism="unknown"
BASE COUNT 710 a 1030 c 529 g 552 t
ORIGIN

Alignment Scores:
Pred. No.: 0.107 Length: 2821
Score: 175.00 Matches: 56
Percent Similarity: 37.36% Conservative: 12
Best Local Similarity: 30.77% Mismatches: 72
Query Match: 17.11% Indels: 42
DB: 6 Gaps: 7

US-10-081-935-4 (1-194) x ARI29836 (1-2821)

QY 25 ProCysAlaLeuAenCysPheValGluAlaLeuGlyAsnAspGlyCysThArgLeuThr 44
|||
Db 1406 CTTGCCACACAGTCTGT-----GAGTCTCTGTGACCGCTGTGC-----CAA 1450

QY 45 AspPheLysCys-----HisCysSerLysProGluLeuProGlyGlnLeuThr 60
|||||
|||

IYVFGSGNGCYDNCVNDGYVSPYTTTIGSIDVRGIRHYFSEOCSSVLATVSGSIV
 TNARITTTDVGEGCSTVHSGSSASTPIAGVIALVLSVSPNTLMDIOGLIYESAVP
 FSLDYGMEKLPSCGRYHYHYEGKLDAYMEANFNKLNQARNSVMTIYNKFE
 SENNGHITDKFNLKAYPDYVNGKLEKERVLTLYFOHAKGSLKINTSSGYTSMKT
 HRLHDKNVCYVMTSTVYKMEPIVGEVTTIDIEDOKITNLGGERYNOLHFGESA
 DLTKPTOSIVPTFTSRSPSPKPGEPVYLPTPSPATISPSVSPQTSSESTPTQPTQ
 PTPQTSSEPTSEPTSEPTSKPTQPTQPTQPTSEPTSEPTSEPTQPPAPQPPAPQ
 PAPPAPQPPAPQPPAPQPPAPQPPAPQPPAPQPPAPQPPAPQPPAPQPPAPQ
 PAPPAPQPPAPQPPAPQPPAPQPPAPQPPAPQPPAPQPPAPQPPAPQPPAPQ
 STEPT
 PAKPT
 ASHTSFKKLLLSILSLISLILILILILILILILILILILILILILILILILILIL
 NMSQTEKATMLRNQOVTSSDSDTDEDPMYERILRS*

BASE COUNT 1055 a 623 c 515 g 921 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 262 Length: 3114
 Score: 169.00 Matches: 32
 Percent Similarity: 61.33% Conservative: 14
 Best Local Similarity: 42.67% Mismatches: 25
 Query Match: 16.52% Indels: 4
 DB: 8 Gaps: 1

US-10-081-935-4 (1-194) x AF093132 (1-3114)

QY 93 ProProvaAspThrAlaAlaProGluProSerGluThraAlaGluProThraAlaGlu 112
 Db 2008 CCACACCTGAGTCATCTGACCAACACCTGACCAACACCTGACCAACACCTGAC 2067
 QY 113 ProThraGluGluProThraAlaGluProThraAlaGluProThraAlaGlu 132
 Db 2068 CCACACCTGAGTCATCTGACCAACACCTGACCAACACCTGACCAACACCTGAC 2127
 QY 133 ProThraGluGluProThraAlaAlaProThraGlyThraGlyGlyGlyAlaProThraGlyThr 152
 Db 2128 CCACACCTGAGTCATCTGACCAACACCTGACCAACACCTGACCAACACCTGAC 2175
 QY 153 GlySerPheThraValThraGlyArgProThraAlaSerThraProAla 167
 Db 2176 CCACACCTGAGTCATCTGACCAACACCTGACCAACACCTGACCAACACCTGAC 2220

RESULT 19
 AC128838/c 256635 bp DNA linear HTG 19-NOV-2002
 LOCUS Rattus norvegicus clone CH230-395E20, *** SEQUENCING IN PROGRESS
 DEFINITION *** 4 unordered pieces.

ACCESSION AC128838
 VERSION 3 GI:25073624
 KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 256635)
 AUTHORS Muzny, D., Matie, M., Metzger, M., Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, B., Baden, B., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, F., Bismalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, M., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Drepper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J.,

TITLE
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M., Hollins, B., Howelle, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovac, C., Kowis, C., Kraft, C. L., Ledow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, B., Louie, L., Louie, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Matlock, E., Mathew, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Mlosoavljevic, A., Miner, G., Ming, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naif, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okunolu, G., Olarnpungson, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Plankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smaiz, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wietzyk, R., Wooten, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhou, X., Zhou, X., Dunn, D., von Weinstock, G., and Gibbs, R. A.
 Direct Submission
 Unpublished
 2 (bases 1 to 256635)
 Morley, K. C.
 Direct Submission
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 256635)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Nov 19, 2002 this sequence version replaced gi:23915267.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KAVH
 Center clone name: CH230-395E20
 ----- Summary Statistics
 Assembly program: Phrap, version 0.990329
 Consensus quality: 199691 bases at least Q40
 Consensus quality: 203586 bases at least Q30
 Consensus quality: 205850 bases at least Q20
 Estimated insert size: 203171; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

Db 496 GCCCAGAGCTGATCAGCGCTGCGCGCATCCCGTCGCGCGCAGCCGCTTC 555
Qy 32 ValGluAlaLeuGlyAsnAspGlyCysThrArgLeuThrAspPheLysCysHisCys--- 50
Db 556 CTCACAGCGCCCGCCGACCATTCGCTCCGCGC---ACCACTTCAAGTCCAGTCCAG 612
Qy 51 SerLysProGluLeuProGlyGlnIleThrProCysValGluGluAlaCysPro----- 68
Db 613 CACACAGCCCAAGATGTTTGGCGCGCTCGAGAGCTGCGTCCAAAGACTGCCAGACT 672
Qy 68 ----- 68
Db 673 GAGTCCAAAAGACATTCAGCGATCTGACAAAGGTAGTGTCTGCGCGCAAGATGC 722
Qy 69 LeuAspAla-ArgIleSerValSerAsnIleValIleAspGlnCysSerLysAlaGlyVa 88
Db 733 GCACAAATACGTGGATGATTCACAAACAAATATGATACCAAGACAGTCTGCACGTCG 792
Qy 88 1ProIleAspIleProProValAspThrAlaAlaProGluProSerGluThrAlaG1 108
Db 793 GCTTCGAGAGTCCCGCTTCACAAACAGCGCGGCGCGCGAA-----AT 837
Qy 108 uProThrAlaGluProThrGluGluProThrAlaGluProThrAlaGluProThr----- 126
Db 838 ACCGTCAACCCGTCAGCTTCATC-CCAGCGCCGACGCTCGACTCTCCACACACAC 896
Qy 127 ----AlaGluProThrHisGluProThrGluGluProThrAlaValProThrGlyThrG1 145
Db 897 CGTCGAGCGCCCGACCGAACCAGTGGCGCGACCTCGCGCGTCCACT----- 948
Qy 145 yGlyGlyValProThrGlyThrGlySerPheThrValThrGlyArgProThrAlaSerTh 165
Db 948 ----- 948
Qy 165 rProAlaGluPheProGlyAlaGlySerAsnValArgAla----- 178
Db 949 -----GCCCGCCGCAACATGCGCGCTGAGTGCAGTGCAGATCGTGT 989
Qy 179 ----SerValGlyGlyIleLeuAlaAlaLeuLeuGlyLeu 190
Db 990 TGGTCCGTGGTGGTGGCTTGGGTAGCCCTTGGGCTG 1029

RESULT 21
AF080221 4291 bp DNA linear PLN 09-AUG-1998
LOCUS Candida albicans mycelial surface antigen precursor (CSA1) gene,
DEFINITION complete cds.
ACCESSION AF080221 GI:3406797
VERSION AF080221
KEYWORDS
SOURCE
ORGANISM
Candida albicans
Candida albicans
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Mitosporic Saccharomycetales; Candida.
REFERENCE
AUTHORS
TITLE
Bourbonnais,Y., Lamare,C. and Deslaurliers,N.
Expression cloning of the Candida albicans CSA1 gene encoding a
mycelial surface antigen with similarity to the Coccidioides
imitis antigen 2
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (24-JUL-1998) Biochemistry, Universite Laval, Cite
Universitaire, Ste-Foy, Que G1K 7P4, Canada
FEATURES
source
1..4291
/organism="Candida albicans"
/mol_type="genomic DNA"
/strain="ATCC32354"
/db_xref="ATCC:32354"
/db_xref="taxon:5476"
515..4126
gene

CDs
/gene="CSA1"
515..4126
/gene="CSA1"
/note="cell wall protein; Csa1p; differential expression
in yeast and mycelial forms; similar to Coccidioides
imitis antigen 2"
/codon_start=1
/transl_table=12
/product="mycelial surface antigen precursor"
/protein_id="AAC29486.1"
/db_xref="GI:3406798"
/translation="MLPSIVISIVLASFVSAESSITEAPTTAEDNPYITPVSAKTA
SINGPADRIYDOLPECAKPCMFONTGVTTPCPYMDTGCLCIMPFGAIGSCIAEKK
ODVVASLIGTSCISVAGWDPYMMIPANVQSSLSAATVASSSEOPVETSEAPAS
SGVESSEOPVETSSSEPAETSSSEPAETSSSEPAETSSSEPAETSSSEPAETSS
TPEDNPYITPVSAKTASINGPADRIYDOLPECAKPCMFONTGVTTPCPYMDTGCLC
PFGAIGSCIAEKKODVVAATSLGTSISVAGWDPYMMIPANVQSSLSAATV
PSSSESVETSSSEPAETSSSEPAETSSSEPAETSSSEPAETSSSEPAETSS
SSSTSTSTFRTASINGPADRIYDOLPECAKPCMFONTGVTTPCPYMDAGLCVMPORF
GAIGCVASCKODIVSVTSLGTSISVAGWDPYMMIPANVQSSLSAATVAPTD
SASERTASQESERTSSSEPAETSSSEPAETSSSEPAETSSSEPAETSSSEPAET
ASINGPADRIYDOLPECAKPCMFONTGVTTPCPYMDTGCLCIMPFGAIGSCIAEKK
GQDVVATSLGTSISVAGWDPYMMIPANVQSSLSAATVASSSEOPVETSEAE
SSQNPASSSQPSETASQEPSETSSSEPAETSSSEPAETSSSEPAETSSSEPAET
PYITPVSAKTASINGPADRIYDOLPECAKPCMFONTGVTTPCPYMDTGCLCIMPFG
ATGSCIAEKKGOEVSVTSLGTSISVAGWDPYMMIPANVQSSLSAATVATSDS
ASERVASSESASQVPOETASQSSANNSVSAAPNSVSAAPNSVSAAPNSVSAAPNS
SSGASVVPQSANNSASAPNSNSSLISGVSAPSVSGSTIAQPSITKSDASIT
GPIITDKVITNESGIVFTSTVITTHVEYCDQTSAAVOSACESSAKESASAS
EQVKVITSVVWCSSIOSIESVKTSAEAKHTEVIAACSELSLSAKESAKETVS
LVEVOKSAVAKORTSLAAVOSASVOLAHAQKSEAVAVQTTAVAEAKADEIST
ETVNIKTIVSSGKETGVSAQTVAANTHSAIVAIAMANTKRAFTNLSLVASVFGLF"
515..565
/gene="CSA1"
/note="putative"
566..4123
/gene="CSA1"
/product="mycelial surface antigen"
/note="putative"
638..943
/gene="CSA1"
/note="Region: hydrophobic Cys-rich domain"
121..1516
/gene="CSA1"
/note="Region: hydrophobic Cys-rich domain"
1721..2026
/note="Region: hydrophobic Cys-rich domain"
/gene="CSA1"
/note="Region: hydrophobic Cys-rich domain"
2207..2512
/gene="CSA1"
/note="Region: hydrophobic Cys-rich domain"
2762..3067
/gene="CSA1"
/note="Region: hydrophobic Cys-rich domain"
4067..4123
/gene="CSA1"
/note="encodes putative GPI-anchoring determinant"
BASE COUNT 1142 a 962 c 835 g 1352 t
ORIGIN

Alignment Scores:
Pred. No.: 0.564 Length: 4291
Score: 165.50 Matches: 49
Percent Similarity: 38.694 Conservative: 28
Best Local Similarity: 24.624 Mismatches: 89
Query Match: 16.184 Indels: 33
DB: 8 Gaps: 6

US-10-081-935-4 (1-194) x AF080221 (1-4291)
Qy 14 GlyLeuAlaSerAlaGluProAspIleProProCysAlaLeuAsnCysPheValGlu 33
Db 656 GGTTCGTCGACAGAAATTATGATCAATTGCCAGAGTGTGCCAAGCATGTATGTTCCA 715

[illegible]

gene
CD5

1175..4231
/gene="WAP1"
1175..4231
/gene="WAP1"
/note="contains predicted signal sequence and GPI addition sequences, indicating localization in the cell wall; contains four cysteine-rich repeats similar to the single repeat of Rbt5p and to Coccidioides immitis proline-rich antigenic protein PRA, Wap1p"

/codon_start=1
/transl_table=12
/product="putative cell wall protein"
/protein_id="AG09792.1"
/db_xref="GI:9963992"

BASE COUNT 1327 a 1009 c 932 g 1524 t

ORIGIN

Alignment Scores:
Pred. No.: 0.624 Length: 4792
Score: 165.50 Matches: 49
Percent Similarity: 38.69% Conservative: 28
Best Local Similarity: 24.62% Mismatches: 89
Query Match: 16.18% Indels: 33
DBs: Gaps: 6

US-10-081-935-4 (1-194) x AF254147 (1-4792)

OY 14 GlyLeuAlaSerAlaGlnLeuProAspIleProProCysAlaLeuAdeCysPheValGlu 33
 |||||
Db 1316 CGTTTCTGTGCACAAATTATGATCATTCGCCAGAGTGCCCAAGCATGTATGTTCCAA 1375
 |||||
OY 34 AlaLeuGlyAsnAspGlyCysThrArgLeuThraAspPheLysCysHisCysSerLvsPro 53
 |||||
Db 1376 AACACTGGGTGACCACCATGTCCA---TACTGGGATCTCGGCTTTGTGTATTATGCCA 1432
 |||||
OY 54 GlueuProGlyGlnIleThrProCysValGluGlnJuaLacCysProLeuAspAlaArgIle 73
 |||||
Db 1433 ACATTGCTGTGTCATTGGTTCTTGTGAATTGCTGAGAAGGTAAAGCCCAAAGACGTTGTT 1492
 |||||
OY 74 SerValSerAniIleValIalAspGlnCysSerLysAlaGly----- 87
 |||||
Db 1493 TCTGCTCAAGATTGGGAACCTTCATTGTTCCGTCGTGTGTGAGATCATCACTGCG 1552
 |||||
OY 88 ---ValProIleAspIle----- 92
 |||||
Db 1553 ATGGGCTGCAAAATGTCCAGACAGCATTTTAAGTCTGCTGCTCCACTGCTGTTGCATGCTCT 1612
 |||||
OY 93 -----ProProValaPrThrThraAlaAlaProGluProSergIuThraAlaGluProThr 110
 |||||
Db 1613 TCTGAACAACAGGTGAACATCTTTCGAACCACTGATCTTCTCAGTCTGTTGAATCT 1672
 |||||
OY 111 AlagluProthrArgGluGluProthrAlaGluProThrAlaGluProthrAlaGluProthr 130
 |||||
Db 1673 TCTAACCTGCTGAACACCTCATCTCGAACCTGCTGAGACCTTCATCATCTCGAACCTGCT 1732

```

Oy      131 HisGluProThrGluLupProThrAlaValProThrGlyThrGlyGlyValProThr 150
        :::::
Db      1733 GAGACTTCATCGAATCATCATCGAACAACATCTTCATCTCACTGCTGAACTCA 1792
        :::::
Oy      151 GlyThrGlySerPheThrValThrGlyArgProThrAlaSerThrProAlaGlu----- 168
        :::::
Db      1793 TCAGAAGAAATCTCTCAATCATCATCTGACGCCCA-----TCACTCTCGAAGATTAACCCA 1846
        :::::
Oy      169 -----PheProGlyAlaGlySerAsnValArgAlaSerValGlyGlyIleAla 184
        :::::
Db      1847 TACACCATCTACCCAGATGTTGCCAAGACT-----GCTTCTATCAATGCTTTTGCT 1897
        :::::

RESULT 23
AC095003/c 238046 bp DNA linear HTG 09-MAY-2003
LOCUS Rattus norvegicus clone CH230-6K12, WORKING DRAFT SEQUENCE, 2
DEFINITION uncloned pieces.
ACCESSION AC095003
VERSION AC095003.7 GI:30467553
KEYWORDS HTG: HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
          Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
          1 (bases 1 to 238046)
REFERENCE Muzny,D,Marie, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J,
          Allen,C, Allen,H, Aisbroke,S, Amin,A, Angulano,D,
          Anyaldebech,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
          Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
          Bissolo,K, Blair,J, Blankenburg,K, Blych,P, Brown,M,
          Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
          Cardenas,V, Carter,K, Cavazos,I, Cesari,H, Center,A,
          Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
          Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
          Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
          Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
          Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Baves,K,
          Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
          Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
          Frazer,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
          Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W,
          Guneratne,P, Haaland,M, Hamill,C, Hamilton,C, Hamilton,K,
          Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
          Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M,
          Hollins,B, Howells,S, Huylk,S, Hume,J, Idlebird,D, Jackson,A,
          Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
          Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovac,C,
          Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J,
          Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
          Lorensuhewa,L, Louisedge,H, Lozada,R,J, Lu,X, Ma,J,
          Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A,
          Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E,
          Mawhney,S, McLeod,M,P, McNeill,T,Z, Meenan,E,
          Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S,
          Morgan,M, Morris,K, Morris,S, Ngunidasa,M, Murphy,M, Nair,L,
          Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S,
          Nwackelameh,O, Okwuonu,G, Olarnpunsagoon,A, Pal,S, Parks,K,
          Pascernek,S, Paul,H, Perez,A, Popovic,D, Primus,E, Pu,L,
          Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L,
          Puato,M, Quito,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R,
          Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,
          Rivers,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J,
          Sanders,W, Savery,G, Scherer,S, Scott,G, Shatsman,S, Shen,H,
          Shetty,J, Shvartsbeyn,A, Sison,L, Sitter,C,D, Sma's,D,
          Sneed,A, Sodergren,E, Song,X,Z, Sorrelle,R, Soes,J,
          Steimle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C,
          Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Umanai,K,
          Vales,R, Vera,V, Villaseana,D, Waldron,L, Walker,B, Wang,J,
          Wang,O, Wang,S, Warren,J, Warren,R, Wei,X, White,F,
          Williams,G, Willson,R, Wleczek,R, Wooden,H, Worley,K,
          Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V,
          Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
Mortley,K,C.
TITLE
JOURNAL
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 238046)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24818271.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GBVH
Center clone name: CH230-6K12
----- Summary Statistics
Assembly program: Atlas!
Consensus quality: 226044 bases at least Q40
Consensus quality: 227924 bases at least Q30
Consensus quality: 229373 bases at least Q20
Estimated insert size: 234613; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 236832: contig of 236832 bp in length
* 236833 236932: gap of unknown length
* 236933 238046: contig of 1114 bp in length.
Location/Qualifiers
1. 238046
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="txon:10116"
/clone="CH230-6K12"
1. 1680
/notes="wgs _end-extension
clone end: T7"
6066 _ 6753
/notes="clone_boundary
clone_end: T7
site: BcORI
end_sequence: BH360832"
228068 .228893
misc_feature
misc_feature
misc_feature

```


Estimated insert size: 223073; sum-of-contigs estimation
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 222215: contig of 222215 bp in length
 * 222216 222215: gap of unknown length
 * 222216 225849: contig of 3534 bp in length.
 Location/Qualifiers
 source
 1. 225849
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-163J19"
 misc_feature
 1. 1302
 /note="wgs_contig"
 misc_feature
 214638..217165
 /note="wgs_contig"
 misc_feature
 220987..223215
 /note="wgs_contig"
 BASE COUNT 60210 a 46663 c 49486 g 61178 t 8312 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 25.2 Length: 225849
 Score: 164.50 Matches: 34
 Percent Similarity: 51.658 Conservative: 13
 Best Local Similarity: 37.364 Mismatches: 43
 Query Match: 16.084 Indels: 1
 Gaps: 2
 US-10-081-935-4 (1-194) x AC098905 (1-225849)
 QY 80 VALAAGGInCySerIysAlaGlyValProIleAspIleProFroValAspThrThra1a 99
 Db 210879 GTAACACCTGTCTCAAAACAGCCAGCCAGCAACCAACCAACCAACCAACCAAC 210938
 QY 100 A1AProGluProSerGluThrAlaGluProThrAlaGluProThrGluGluProThrAla 119
 Db 210939 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 210998
 QY 120 GIUProThAlaGluProThAlaGluProThAlaGluProThAlaGluProThAla 139
 Db 210999 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 211058
 QY 140 ValProThGlyThrGlyGlyValProThGlyThrGlySerPheThrValThrGly 159
 Db 211059 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 211118
 QY 160 ArgProThAlaSerThrProAlaGluPhePro 170
 Db 211119 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 211148
 RESULT 25
 DDU20608 4675 bp DNA linear INV 13-MAY-1997
 LOCUS DDU20608
 DEFINITION Dictyostelium discoideum unknown spore germination-specific
 protein-like protein, orf1, orf2 and orf3 genes, complete cda.
 ACCSSION U20608
 VERSION U20608.1 GI:676858
 KEYWORDS
 ORGANISM Dictyostelium discoideum
 SOURCE Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 REFERENCE 1 (bases 1 to 4675)

AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 CDS
 Giorda, R., Ohmachi, T., Shaw, D. R. and Ennis, H. L.
 A shared internal threonine-glutamine-acid-threonine-proline repeat
 defines a family of Dictyostelium discoideum spore germination
 specific proteins
 Biochemistry 29 (31), 7264-7269 (1990)
 91002566
 1976383
 2 (bases 1 to 4675)
 Ennis, H. L.
 Direct Submission
 Submitted (02-FEB-1995) Herbert L. Ennis, Roche Institute of
 Molecular Biology, 340 Kingstand Street, Nutley, NJ 07110, USA
 Location/Qualifiers
 1. 4675
 /organism="Dictyostelium discoideum"
 /mol_type="genomic DNA"
 /strain="Ax3"
 /db_xref="taxon:44689"
 /clone="370G3"
 /cell_type="amoeba"
 join(189..312,392..1110)
 /note="orf1"
 /codon_start=1
 /evidence=not experimental
 /product="unknown"
 /protein_id="AAB54082.1"
 /db_xref="GI:2081632"
 /translation="MEYSLFLKVTGENVEIDFTINQNGHNPVKVTFTEYNNKF
 PNLNSVNPVNLVESDMCTFSNNFKYTSYLQPDTSVVGFGFGSGDNNYKN
 NICEFTPOATYINFTENNVDYKTFNPBELRKLTKKKKKLKKSKYFPLNIN
 INNFIILIAVEFSPIKVESYNILTVSPFTLITINSSVGINAOFILTVSSVSEF
 LISGDKNYTOAVIDVNTPSFRDENGMHLVDSGISSDPIYFKLPSIMYKI
 LKNL"
 313..391
 join(448..1081,1189..3926)
 /note="orf3: similar to Cella, encoded by Genbank Accession
 Number M33661, and Cella, encoded by Genbank Accession
 Number M33662; contains TETP repeat."
 /codon_start=1
 /evidence=not experimental
 /product="unknown"
 /protein_id="AAB54083.1"
 /db_xref="GI:2081633"
 /translation="MFCFSFNNFKYTSYLQPDTSVVGFGFGSGDNNYKN
 CEFTPOATYINFTENNVDYKTFNPBELRKLTKKKKKLKKSKYFPLNIN
 NFIIILIAVEFSPIKVESYNILTVSPFTLITINSSVGINAOFILTVSSVSEF
 LISGDKNYTOAVIDVNTPSFRDENGMHLVDSGISSDPIYFKLPSIMYKI
 LKNL"
 1110
 /note="orf2"
 /codon_start=1
 /evidence=not experimental
 /product="unknown"
 /protein_id="AAB54081.1"
 /db_xref="GI:2081631"
 /translation="MFCFSFNNFKYTSYLQPDTSVVGFGFGSGDNNYKN
 CEFTPOATYINFTENNVDYKTFNPBELRKLTKKKKKLKKSKYFPLNIN
 NFIIILIAVEFSPIKVESYNILTVSPFTLITINSSVGINAOFILTVSSVSEF
 LISGDKNYTOAVIDVNTPSFRDENGMHLVDSGISSDPIYFKLPSIMYKI
 LKNL"


```
mol(d). inositol 5-phosphatase 2"
/protein_id="AA052441.1"
/db_xref="GI:28829951"
/translation="MCGNKHQACQACAVRSKDFEGLFECNCHTRVEGFTQEV
EPEENGCKYKTSYKRVWTOEVLDAWOKLRKQKRYIYLQWLKIQVTSYLSKIDGY
NQQLYOVTELMWFLKIDKFSQDEGSSSTSSSTSTSTFTGIINYPPTLVSICVA
CYLLRPAVTSIDFISLAOSGKIASFTNETTSSSSNSTPINFARISDDCKYANL
LSSSNPLLLRFCETELKPKSLINLNOLYNLILPFLFKKSPNLSLTKYENIL
LPSNSDLPLVLSIIMLILKSKFNLFDROYNPNSSGSSNNNNNTLNDNNNNNNNN
NNEYDNNNDKYNKENSIERLEFLKLNNNNFENSENSISSEERKEEYKRNEL
ENWLGVYIYNLPKNEFLSKODTWSLIDQPGIPKLPSSIRRLITDQIHSTKXDSLSD
KYQNLNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
SKSKSSSSSSSSSSSSADPTILLPKVNNIRAKSINRSGRTNLPFPQKSVFSKGT
GKRG"
```

6218. .9748

/note="GeneID exon scores (in order of location ranges):
356.10 - GSCJ_ID dd_03095"

/codon_start=1

/product="similar to Dictyostelium discoideum (Slime
mold). kinase responsive to stress 1-like kinase"

/protein_id="AA052442.1"

/db_xref="GI:28829952"

/translation="MSNLKSNNGNGQESSFFQKVKSPSTONLNSFSSNSNN
NLNSGSEVKKDTTNSPOLPENYTPPPPPHOIRNSSIEGGEFLNNNSDNNNN
NN
TDDSTTETETGMSPVQTLNSIDSEQOHODLSSSVNNNNNNNNNNNNNNNNNN
NN
PGLKISPPNSDQPLKEKKKKKKFKTPPEIKHHHKKSSSSSTTTPSTTSSTSP
SSSSLAISSPNTAATTNNKSHKTKTSTFDINTEISVYNNVHKHNVDFDKWGN
DTLDEKLDGAGVGVYKGTNLDLQFTLAIKVIEMKESVSLQNEINLKQKSPNI
VSYFSLQTESHWILDLFCALGIRDIIESTEKLENAQISFVVKNTLKQIYLSQ
NIHRDVKANVLLSECDVKIADFGVSEKNGALDOSEKMITPLMAAPYILKQY
DYKADIMSLGITIEADGLPHIDLPMRAMKWNMPPEPAPKPKSLINPFLA
RLVLDPEKRAPIDLCHPLKDRGPRVLDVNLQLEKIKKKIKDLKKQKHQTS
QSSSSSPSPRNATVNGDIDGSLSTISSTPSDELNCNNSLNLAAVSKML
SKGVYTRDRDQEEEDSKLNNDQEDQDEDDDEDDDEDDDEDDDEDDDEDDDE
KKKGGGNGGVTSQDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE
NKTSTKSNKKKKNNKNNLSTIGSGGNNLHVAHPKPLPTPPGISMTNEKOLET
LETTYIDSNQIKVINDIKNEIKOLESSIIKININLOQOLSPILLALEEKONHTT
PROMOSKSLATNLNEKULSSSPSSPLTNSVNSLTTTSTTTTTPULSRGSSRFS
GSISSSSSPRNASAIMAVNNSSTTTTNSSSSSNGGSGVDISPTMGASPSIM
KRFTTSSSSSPSSSDGFAFNSSSTTTTNSSSSSSSSSSSSSSSSSSSSSSS
NSSGSGSSPSLSTSSSTNSNNSVNNIKK"

join(10002. .10168.10786. .12082)

/note="GeneID exon scores (in order of location ranges):
0.26, 60.31 - GSCJ_ID dd_00942"

/codon_start=1

/product="hypothetical protein"

/protein_id="AA052443.1"

/db_xref="GI:28829953"

/translation="MESIFYSYLNNYDCGNDEEFLNCLLITSLNNLNTFENLLY
CYNNLKRKTLNFKVNLIIEDNNGETGFTPLNRPVLPYGVREVDVDCSSITIG
DHWIPDCKSYFHNPNQOKLNRNINSLTITFGNSGNOLASNGDPMPLKLYKT
LSLQGSFOOTORNELPOSTLEPLDSEYQGIILNESIPKSVYTLTAFFETIRKHF
TKSISKGVTSKIGCYFNQITKDMIPINWTSLEIDNSKIEIBNSPISATKXIRIG
SIELSDNLIPNINKLSYGDVHSCFLPKSIESIKITSFNLKISFNPNLNLSTK
IEKLSDSVLEBGFNPNTLRLIYSCPNLKSNGIPNSIKKLKKEPNHPIANGPLPS
SIEVLKPNYDIPKNGDLPSSLNVLKLDHCTISFIODALPNNKLKTYFHRELRKL
PNLWPTSPPEFYISEASFPNLSIDLDFETHILK"

join(12253. .12311.12539. .13007)

/note="GeneID exon scores (in order of location ranges):
-5.63, 15.15 - GSCJ_ID dd_00941"

/codon_start=1

/product="hypothetical protein"

/protein_id="AA052444.1"

/db_xref="GI:28829954"

/translation="MDNITISYLNNYNNKNNNEEKEEKKLANDLQKLANDQOFTLN
NHIIQNLKNTFTTNSNDVIYNNHSITDNNNSNDIINNRKNIADCSILLIINKNLN
HHIITKNNNSNNLNLILKPPDPLNNKIRITQOQIDQITDNNYKLFKVMGNIIFRDO
ILFLRLIYNNHNSI"

join(13144. .14052.15655. .15766)

/note="GeneID exon scores (in order of location ranges):
36.56, 7.78 - GSCJ_ID dd_00940"

CDS

```
/codon_start=1
/product="hypothetical protein"
/protein_id="AA052445.1"
/db_xref="GI:28829955"
/translation="MRCSDIILVDHNIPTCKSIIPSHNQOLTRNIIIGNSLTSIDFG
KPNQSLSDKNGPMLPRLTKLSIFGKSPQSHIHREBIPSLSLDLPYKAVIEG
SIPKSVTLTHADYFWSKGAKLDSITPGGTSTLSPFQOITSEKMSIHNNTVNS
FNSGISLSPISITKLHGSNONPPRVQVATNNLSYTTSSDIKILPKTIQKLI
MPSPPSQLDNHSFYLRKULSINFDAKQVDLNVFAPLIRANGLIQALLSDHGLVNNLIT
FNQLOD"
```

join(15826. .16011.16072. .17106)

/note="GeneID exon scores (in order of location ranges):
1.92, 31.02 - GSCJ_ID dd_00422"

/codon_start=1

/product="hypothetical protein"

/protein_id="AA052446.1"

/db_xref="GI:28829956"

/translation="MTQLFRMKRIAFDCSIIIIKNNLHHITIKNNLHNTTNSNN
KNNLNOVITLIDINKGEYLSGKVFLEYIIIIISIDIGGHRISYKTRN
YRNSVINSDENNHQQNISNCCPMPSTAEKIEKMSDIIIVDHWIPDCKSIKFTY
FDQQLRNLIIIGNSLTSINFSEFNQSLDSTKGPMLPRLKSLTFGKSPQSHIHREB
PSSLTLTFIPYKGVIOYGSIPKSVTLTHADYFWEKHDIIYKFDLTSVPGTSLFED
FYNGTIEKGMIPHYVTSLSKFRMGDKLEPNSLPISIKTLGCRNRTIFQKVPPTI
TNLSITVTSIDIGILPTEKIKIIGSNFTQDLSHFHIFENKLSLMAIGVDLTGV
EESPITKTLTLVL"

join(18782. .18911.19094. .20841)

/note="GeneID exon scores (in order of location ranges):
0.11, 70.39 - GSCJ_ID dd_00420"

/pseudo

/codon_start=1

join(21335. .21528.21616. .21928)

/note="GeneID exon scores (in order of location ranges):
-5.52, 10.92 - GSCJ_ID dd_00435"

/codon_start=1

/product="hypothetical protein"

/protein_id="AA052447.1"

/db_xref="GI:28829957"

/translation="MDSITFYLYNNYNNKNNNEELLNCLLITSLNNLNTFNOVODYC
NSINNYKNNYONLQKKKCKLOEEDBEETHLNDLKGINDQOFTLNHNNLNTFTN
NSNDEITYTHSITDNNNNNDITDNRKRIAFDCSIIIIKNNLHHITIKNNNNNN
NLQYLK"

complete(22154. .22339)

/note="GeneID exon scores (in order of location ranges):
0.40 - GSCJ_ID dd_00433"

/codon_start=1

/product="hypothetical protein"

/protein_id="AA052448.1"

CDS

Alignment Scores:

Pred. No.:	14.8	Length:	110000
Score:	163.50	Matches:	54
Percent Similarity:	37.71%	Conservative:	12
Best Local Similarity:	30.86%	Mismatches:	82
Query Match:	15.98%	Indels:	27
DB:	3	Gaps:	4

US-10-081-935-4 (1-194) x AC116957_0 (1-110000)

QY	22	AspIleProProCyAlaIeuuBnCySphreValaIguAlaIeuGlyAAsnAP-----	38
DB	48800	GATGTTCCAGTATTATACAGGTTCTGTATGAATTTGTACTTAAAGTGCGTATTATAGCTTA	48859
QY	39	-----GlyCyethrArgIeuUthrAphrPheCySnhIaCySerIyBProGluIeu	55
DB	48860	ACAGCTAAAGATTAAATCATCTACTGTCGACATTCATCAATTATTAATCACTATAA	48919
QY	56	ProGlyGlnIethrProCyValaGluGluIaCySProIeu-----	69
DB	48920	CCATTACAAAGTAAAAAAGCTTGCGGCGTCTGCTATTATTTATTAAGCGCGATCTCTGA	48979
QY	70	-----AspAlaArgIeserValser	76

```

Db      48980 ATGGATCAACAGTACTTACTTATTATTTAGTAAATGATATAATCAATATCTTCAAT 49039
Oy      77 AsnIleValaIaspInCysSerIyAlaGlyValProIleaspIleProValaIasp 96
Db      49040 GAATATTATTATTATTCATATCAAAAAGTATTATTATTATTTATTCAGATCCACCAACCCCA 49099
Oy      97 Thr---ThraIAlaIaProGluProSeGluThrAlaGluProThrAlaGluProThrGlu 115
Db      49100 ACTCAACACACCAACACCACTGAAACACCAACCAACCACTCAACACCACTCAAA 49159
Oy      116 GluProThrAlaGluProThrAlaGluProThrAlaGluProThrAlaGluProThrGlu 135
Db      49160 ACACCACTCAACACCACTCAACCTCAACCTCAACCTCAACCTCAACCTCAAC 49219
Oy      136 GluProThrAlaValProThrglyThrglyGlyValProThrglyThrglySerPhe 155
Db      49220 ACACCACTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCTCA 49279
Oy      156 ThrValThrglyThrglyProThraIaSerThraProIaGluPhePro 170
Db      49280 ACACCAACCGAAACACCACT--GAAACACCACTGAAACACCA 49321

RESULT 27
AC119773 222896 bp DNA linear HTG 19-NOV-2002
LOCUS Rattus norvegicus clone CH230-150P22, *** SEQUENCING IN PROGRESS
DEFINITION
AC119773
AC119773.6 GI:25095179
VERSION HTG: HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Sciuromorphi; Muridae; Murinae;
Rattus.
1 (bases 1 to 222896)
1 Muzny,D,Marie, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J,
Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angiano,D,
Alyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
Biswal,K, Blair,D, Blankenburg,K, Blyth,P, Brown,M,
Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Center,A,
Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
Devilla,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
Deigado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Escotio,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,M,
Gunsberg,P, Haaland,M, Hamill,C, Hamilton,K, Hamilton,X,
Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M,
Hollins,B, Howells,S, Huliy,S, Hume,J, Idlebird,D, Jackson,A,
Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
Karpachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovat,C,
Kovits,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J,
Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
Lorenshewa,L, Louissege,H, Lozada,R,J, Lu,X, Ma,J,
Maheshwari,M, Mahindartne,M, Mahmoud,M, Malloy,K, Mangum,A,
Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E,
Mawhney,S, McLeod,M,P, McNeill,T,Z, Meenen,E,
Milosavljevic,A, Miner,G, Minje,E, Montemayor,J, Moore,S,
Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L,
Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S,
Nwokilemeh,O, Okwou,G, Olampunsgoon,A, Pal,S, Parks,K,
Paternek,S, Paul,H, Perez,A, Popovic,D, Primus,E, Pu,L,L,
Plopper,F, Poindecker,A, Popovic,D, Primus,E, Pu,L,L,
Plazo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R,
Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,R, Riggs,F,
Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J,

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Sander, M., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soer, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanli, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willison, R., Wleczek, R., Wodden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G., and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 222896)
Worley, K.C.
Direct Submission
Submitted (02-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 222896)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:22856367.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GVPs
Center clone name: CH230-150P22
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 206295 bases at least Q40
Consensus quality: 208887 bases at least Q20
Estimated insert size: 209030; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/centbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 7 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1
130230: contig of 130230 bp in length
130231 130330: gap of unknown length
130331 150793: contig of 20463 bp in length
150794 150893: gap of unknown length
150894 217180: contig of 66287 bp in length
217181 217280: gap of unknown length
217281 218456: contig of 1176 bp in length
218457 218556: gap of unknown length

```

[illegible]

REFERENCE 2 (bases 1 to 4284)
AUTHORS Giorda, R., Ohmachi, T., Shaw, D.R. and Ennis, H.L.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1995) Herbert L. Ennis, Roche Institute of
Molecular Biology, 340 Kingsland Street, Nutley, NJ 07110, USA

FEATURES
source
location/Qualifiers
1..4284
 /organism="Dictyostelium discoideum"
 /mol_type="genomic DNA"
 /strain="Ax3"
 /db_xref="taxon:44689"
 /clone="270-P"
 /cell_type="amoeba"
join(4..1250,1733..3117)
 /note=ori3; similar to Cella, encoded by GenBank Accession
 Number M33861, and Cella, encoded by GenBank Accession
 Number M33862; contains TETP repeat"
 /codon_start=2
 /evidence=not experimental
 /product="unknown"
 /protein_id="AAB54079.1"
 /db_xref="GI:2081630"
 /translation="NSLPTETILNLCFOTPSKRIKSHOLIKHPFENDENOFYNKKE
YFPLKENERDSYIEIHNTESIGSNTCSINEIRFNLYLIGTYENQYIKITLHE
KYGISLSHLNKKFIYLLFLPLMTLVNLNHNVOFTGIORDNFIISTPS
NP1KKSPQSSDNYQYSEGSQSYESSSSSESSSESSSESSSESSSESSSE
NTVFISQKIDQCLTVLNSQLEFIDKGRDSMVLEGNAEOTFSIREKGMATC
LSEBHFHSFKLGRNLANKDGRDCTFNLI1GTQIDQANLVFSRSPNDYIOSDET
RFISTKRGAGSQSQOFFYFHSLSLNMKNLEIKVNNNNNNNNNNNNNNNNNN
NNNNNDLSISINIDFGNIIISODKISQFPKSKIKKSRPALTLYLFLNFKKI
IHVEFIFSLFMASLVNRLIPIHNTSOSDNLFSHLNGLPGLSKNPLNQTSSS
ISFSSSIOPLSESSSOSLTSSESSOSIGSSSSSSONPSSTPESPTPTETETPT
TEMPPT
TEPT
APPTPPYDPSPKYIISKERLMAKTSKEKELKLNHTRGTEORLRIIDGDTLNRI
FYKSCICLKGDPRFTFNKYGGFSGAODECIFTYILNADGTYNFKNKONTYLSDSG
EEFFGRSLSEBAFRLDEAF"
<1..1267
 /note="orf1"
 /codon_start=2
 /evidence=not experimental
 /product="unknown"
 /protein_id="AAB54076.1"
 /db_xref="GI:684938"
 /translation="NSLPTETILNLCFOTPSKRIKSHOLIKHPFENDENOFYNKKE
YFPLKENERDSYIEIHNTESIGSNTCSINEIRFNLYLIGTYENQYIKITLHE
KYGISLSHLNKKFIYLLFLPLMTLVNLNHNVOFTGIORDNFIISTPS
NP1KKSPQSSDNYQYSEGSQSYESSSSSESSSESSSESSSESSSESSSE
NTVFISQKIDQCLTVLNSQLEFIDKGRDSMVLEGNAEOTFSIREKGMATC
LSEBHFHSFKLGRNLANKDGRDCTFNLI1GTQIDQANLVFSRSPNDYIOSDET
RFISTKRGAGSQSQOFFYFHSLSLGNPLD
1251..1732
1735..3117
 /note="internal repeat protein; putative spore germination
specific protein, similar to Cella, encoded by GenBank
Accession Number M33861, and Cella, encoded by GenBank
Accession Number M33862; contains TETP repeat"
 /codon_start=1
 /evidence=not experimental
 /product="unknown"
 /protein_id="AAB54078.1"
 /db_xref="GI:684940"
 /translation="MKGLIKVNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNDLSIS
INDPNNTIISODKISQFPKSKIKKSRPALTLYLFLNFKKIYHVEFIFSLP
MALVYNRLIPIHNTSOSDNLFSHLNGLPGLSKNPLNQTSSSISFESLOPL
ESSENQSLTSSESSOSLGSSESSONPSSTPESPTPTPTPTPTPTPTPTPTPT
PT
PT
PYIISERLMAKTSKEKELKLNHTRGTEORLRIIDGDTLNRIIDFPGKSI
CLDKDPRFTFNKYGGFSGAODECIFTYILNADGTYNFKNKONTYLSDSGGEFFGRKSGS
EEFFGRSLSEBAFRLDEAF"


```

Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: GDH2
Center clone name: CH230-9E3
-----
Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 22969 bases at least Q40
Consensus quality: 225848 bases at least Q30
Consensus quality: 227780 bases at least Q20
Estimated insert size: 241532; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
NOTE: This is a "working draft" sequence. It currently
consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that have
provided by the submitter.
This sequence will be replaced
by the finished sequence as soon as it is available and
the accession number will be preserved.
1 252801: contig of 252801 bp in length.
Location/Qualifiers
1..252801
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-9E3"
1..1620
misc_feature
1..1620
/notes="wgs contig"
misc_feature
251753..252801
/notes="wgs contig"
BASE COUNT 63531 a 49651 c 49432 g 66454 t 23733 others
ORIGIN
-----
Alignment Scores:
Pred. No.: 41.8 length: 252801
Score: 161.50 Matches: 44
Percent Similarity: 44.858 Conservative: 17
Best Local Similarity: 32.354 Mismatches: 64
Query Match: 15.794 Indels: 11
Gaps: 3
-----
US-10-081-935-4 (1-194) x AC095211 (1-252801)
OY 51 SerTysProGluLeuProGlyInIleThrProCySValGluGluAlaCysProLeuAsp 70
Db 185533 AGCCAAACCAATCAGCCCAACCAACCAACCAACCAATCAGCC-----TACCA 185480
OY 71 AlaATGILeserValSerAnIleValIaAspGInCysSerTys-----AlaGlyVal 88
Db 185479 GCAACCAAGCAACCAAGCAACCAACCAACCAACCAACCAACCAACCA 185420
OY 89 ProIleAspIleProProValAspThrThrAlaIaProGluProSerGluThrIaGlu 108
Db 185419 CCGACTGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 185360
OY 109 ProThrIaGluProThrGluGluProThrAlaGluProThrAlaGluProThrAlaGlu 128
Db 185359 CCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 185300
OY 129 ProThrIaGluProThrGluGluProThrAlaGluProThrAlaGluProThrAlaGlu 148
Db 185239 CCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 185240
OY 149 ProThrGlyThrGlySerPheThrValThrGlyArgProThrAlaSerThr----- 165
Db 185239 CCAACCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 185180
OY 166 -----ProAlaGluPheProGlyAlaGlySerAsnValArg 177

```

```

Db 185179 CCGACCACTAACCGAAAAAACCCTCGGCTTCGACCACTGAACCTTACA 185132
RESULT 30
AE012556/c
LOCUS
DEFINITION
Xylella fastidiosa Temecual, section 4 of 9 of the complete
genome.
ACCESSION
AE012556 AE009442
VERSION
AE012556.1 GI:28056730
KEYWORDS
SOURCE
ORGANISM
Xylella fastidiosa Temecual
Xylella fastidiosa Temecual
Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
Xanthomonadaceae; Xylella.
REFERENCE
1 (bases 1 to 300029)
Van Sluys,M.A., de Oliveira,M.C., Monteiro-Vitorello,C.B.,
Miyaki,C.Y., Furlan,L.R., Camargo,L.E.A., da Silva,A.C.R.,
Moon,D.H., Takita,M.A., Lemos,E.G.M., Machado,M.A., Ferro,M.I.T.,
da Silva,F.R., Goldman,M.H.S., Goldmann,G.H., Lemos,M.V.F.,
El-Dorry,H., Tsai,S.M., Carreir,H., Carraro,D.M., de Oliveira,R.C.,
Nunes,L.R., Siqueira,M.J., Coutinho,L.L., Kimura,E.T., Ferro,E.S.,
Harakava,R., Kuramae,E.E., Martino,C.L., Gigliotti,E., Abreu,I.L.,
Alves,L.M.C., do Amaral,A.M., Baia,G.S., Bianco,S.R., Brito,M.S.,
Camavari,F.S., Celestino,A.V., da Cunha,A.F., Fenille,R.C.,
Ferreiro,J.A., Formighieri,E.F., Kishi,L.T., Leoni,S.G.,
Oliveira,A.R., Rosa Jr.,V.E., Sasaki,F.T., Sena,J.A.D., de
Souza,A.A., Truffi,D., Tsukumo,F., Yanai,G.M., Zeros,L.G.,
Civerolo,E.L., Simpson,A.J.G., Almeida Jr.,N.F., Setubal,J.C. and
Kitajima,J.P.
Comparative Analyses of the Complete Genome Sequences of Pierce's
Disease and Citrus Variegated Chlorosis Strains of Xylella
fastidiosa
J. Bacteriol. 185 (3), 1018-1026 (2003)
JOURNAL
PUBMED
1253478
REFERENCE
2 (bases 1 to 300029)
Van Sluys,M.A., de Oliveira,M.C., Monteiro-Vitorello,C.B.,
Miyaki,C.Y., Furlan,L.R., Camargo,L.E.A., da Silva,A.C.R.,
Moon,D.H., Takita,M.A., Lemos,E.G.M., Machado,M.A., Ferro,M.I.T.,
da Silva,F.R., Goldman,M.H.S., Goldmann,G.H., Lemos,M.V.F.,
El-Dorry,H., Tsai,S.M., Carreir,H., Carraro,D.M., de Oliveira,R.C.,
Nunes,L.R., Siqueira,M.J., Coutinho,L.L., Kimura,E.T., Ferro,E.S.,
Harakava,R., Kuramae,E.E., Martino,C.L., Gigliotti,E., Abreu,I.L.,
Alves,L.M.C., do Amaral,A.M., Baia,G.S., Bianco,S.R., Brito,M.S.,
Camavari,F.S., Celestino,A.V., da Cunha,A.F., Fenille,R.C.,
Ferreiro,J.A., Formighieri,E.F., Kishi,L.T., Leoni,S.G.,
Oliveira,A.R., Rosa Jr.,V.E., Sasaki,F.T., Sena,J.A.D., de
Souza,A.A., Truffi,D., Tsukumo,F., Yanai,G.M., Zeros,L.G.,
Civerolo,E.L., Simpson,A.J.G., Almeida Jr.,N.F., Setubal,J.C. and
Kitajima,J.P.
Direct Submission
Submitted (17-DEC-2001) Instituto de Biociencias, Universidade de
Sao Paulo, Rua do Matao, 277, Sao Paulo, SP 05508-900, Brazil
FEATURES
source
1..300029
/organism="Xylella fastidiosa Temecual"
/mol_type="genomic DNA"
/strain="Temecual"
/db_xref="taxon:183190"
/notes="Pierce's disease strain"
113..601
/gene="xpgC"
/locus_tag="PD0734"
/locus_tag="PD0734"
/notes="synonym: pefG"
113..601
/gene="xpgC"
/locus_tag="PD0734"
/locus_tag="PD0734"
/notes="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="general secretory pathway protein G precursor"
/protein_id="AA028603.1"

```


gene /db_xref="GI:28056731"
/translation="MRSAFVFPFYFLIRPEVAVMNRRTSCSPAQRQAGMLLEII
VLVIGVAVFVSGRYLGGADKALVKSQIQLAGKIENYQLDGLKPLGSLNDLVN
PFGVSNMIGPYAKAAELNDPMGHPIQYRVPEGRPFDMISLKGDKPGSSSYDADIK
YE"
607. 1113
/gene="xpsH"
/locus_tag="PD0735"
607. 1113
/gene="xpsH"
/locus_tag="PD0735"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="general secretory pathway protein H precursor"
/protein_id="AA028604.1"
/db_xref="GI:28056732"
/translation="MLPCSAFPFRFQSSGAPVDAQGVSLLEMLVIALIALISLGLVA
VLSGMDVQLRNSKQLAELKRTVALATRAFRMIDPQRRKQGNHOGKLP
GALQIRFSGAQLQTSSEGVFIQFPDGAASSGRIDLEMKRAYRIDVGWITGEVRSG
PLRTSASL"
110. 1508
/gene="xpsI"
/locus_tag="PD0736"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="general secretory pathway protein I precursor"
/protein_id="AA028605.1"
/db_xref="GI:28056733"
/translation="MKCGRYSLLEVIYAFALIALSLILGLTGSTRVRRADLAS
QATLVQSLAQGVAMPRLVHQGGQFEGFRRMILETRPENSGSPGAOLLELI
LSVRWDEPDQSIQWRTLRFVRLQPPDGL"
1505. 2131
/gene="xpsJ"
/locus_tag="PD0737"
/note="synonym: pefJ"
1505. 2131
/gene="xpsJ"
/locus_tag="PD0737"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="general secretory pathway protein J precursor"
/protein_id="AA028606.1"
/db_xref="GI:28056734"
/translation="MRMAGFTLMEVLLATVLLASGLTLPFTLHSAISVRSSEALA
SRMRRAVELFLRRMGLPLMGGEDRRHMLFVGRRMRFAADPDLVLRGCG
PYHDISVSGVGHQRLALTMLOSQETKDGISVSELLADGLHVVHYSKMGDOS
RHQDGPVATWPSDDLPSVAVOIKODIAPWPLVVLQNSDGDG"
2131. 2976
/gene="pefK"
/locus_tag="PD0738"
2131. 2976
/gene="pefK"
/locus_tag="PD0738"
/locus_tag="PD0738"
/locus_tag="PD0738"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="general secretory pathway protein K"
/protein_id="AA028607.1"
/db_xref="GI:28056735"
/translation="MRRAGAAALIVLMLVLAAMIGAFALSAVHSLQCAVLVSSA
QAOEVAGLEAVSRLODSQVVRWIPQGRPYSMYADATLEMRMIDSGKVDLV
ADQLLTALIQALGVDAARAARLAGAIIWMRTDHLVSPGAEBSDYAAAGLPGAGD

gene NPFQSIARELVLGMDTDLYRLMLPNITISYGNLOPEVSAPVPLTALGQNVAAVLA
KPOQSPSPVSVSKOMSGTSTIESKRLACREAVLRAVVRMGGRIPGATYITLHMBE
GIAMQ"
2949. 4103
/gene="pefL"
/locus_tag="PD0739"
2949. 4103
/gene="pefL"
/locus_tag="PD0739"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="general secretory pathway protein L"
/protein_id="AA028608.1"
/db_xref="GI:28056736"
/translation="MGRGYRMAKMSKMSLRFVGSIAQVCGFMHMMQSLLAUPLM
RCRVOMGLSERLLSVQADGHLVYQCCDVEPLVQLPMPITPOELSGMLPLKQL
PRHWLPLATSALCRPLRLPAGASORLHDVDFEIDBOTFEQVQVYDVRLKRGDD
ADLETLLVVPRLVNDGPRGVKEMATTAGIDVADMOGOLNVNLLPQVRKSDP
MRMRNLAAVALLVATGALLNNRIRIVABOLRLQVQOQASARSARVAERORLVT
IEGVRLDERRTLRLPVIEVWELSRRLPNGTYLEKFALEDGOLHMGNSDSSLVR
LLEGSKLMRTSPSLSVLSQNSVSVNRVITAEVLGVVPKKAADGAOR"
4084. 4716
/gene="xpsM"
/locus_tag="PD0740"
4084. 4716
/gene="xpsM"
/locus_tag="PD0740"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="general secretion pathway protein"
/protein_id="AA028609.1"
/db_xref="GI:28056737"
/translation="MARHVNYSRMLALGLFIKLSVLYLLVHPMFTPRMYSVQOQI
QVIRBELRITKQLQGANLAEQLRLIOLVLRBPFLPSSSVLSAGLVKLEGV
NDASPNRSCATITNRSPLQBDMDHDFRYRAVVOVRLCGTFELALVYNLESGPRLFV
NMLVNAQRYPPLSPNSGHDVAFELAGFLSPGQHVPSLPORSSNAN"
4706. 5473
/gene="xpsN"
/locus_tag="PD0741"
4706. 5473
/gene="xpsN"
/locus_tag="PD0741"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="XpsN protein"
/protein_id="AA028610.1"
/db_xref="GI:28056738"
/translation="MRIDAFSVTWLFAVAGAAALVCMALAGMGRSVOLLRDSDQL
IRKQLSPASAEHGPFRVGEIARPVESDRKHPPLVSGEITSVEDRLTG
VLTTRPKATYLTNEONOSLRIGCDVAVVSLIOPSATIEGAGRTLELOF
NGDANKVANNPSENNVPTNVEHVVAVSSSAALALVXPSTSAVSOSGRPQCGQPA
VTGPQAPKRPIDQAIRAIRERIEARRRQLOOQOQK"
5518. 7845

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
48.9	300029	39	4	49	1	1
Score:	161.50					
Percent Similarity:	46.24%					
Best Local Similarity:	41.94%					
Query Match:	15.79%					

US-10-081-935-4 (1-194) x AE012556 (1-300029)

Qy 79 ValValaspGInGysSerlysalaglyValprolleAspIleProProValAspThrThr 98
|||:::|||||
Db 252627 GTCTTAGAGCCCTGTGTGATTAAGTTCGATAGATACACCTACACGACACCCACA 252568

Db 1124 TTTTATT 1116

RESULT 32
AC109525 0/c

WPCOMMENT
Sequence split into 5 fragments LOCUS AC109525 Accession AC109525

Fragment Name	Begin	End
AC109525_0	1	110000
AC109525_1	100001	210000
AC109525_2	200001	310000
AC109525_3	300001	410000
AC109525_4	400001	478059

LOCUS AC109525 478059 bp DNA linear HTG 11-OCT-2002
DEFINITION Rattus norvegicus clone CH230-255D23, *** SEQUENCING IN PROGRESS

ACCESSION AC109525 6 GI:23603814
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE
AUTHORS 1 (bases 1 to 478059)
Mazny,D,Warie, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J.,

Allen,C, Allen,H, Albrooks,S, Amin,A, Angiano,D, Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Bisswaio,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cadenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gbregoregis,E, Geer,K, Gill,R, Grady,M, Guerra,M, Guevara,W, Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M, Hollins,B, Howells,S, Huliy,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorenshewa,L, Louisedge,H, Lozada,R,J, Lu,X, Ma,J, Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangun,A, Mangun,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Mahliney,S, McLeod,M,P, McNeill,T,Z, Meenen,E, Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,S, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwokilemeh,O, Okunnu,G, Olarunsaagun,A, Pal,S, Parks,K, Patelrak,S, Paul,H, Perez,A, Perez,A, Piamkoch,C, Plopper,F, Poindexter,A, Popovic,D, Primus,B, Pu,L,L, Puzo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M, Reigh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savery,G, Scherer,S, Scott,G, Shatman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Snajls,D, Sneed,J, Sodergren,E, Song,X-Z, Sorille,R, Sosa,J, Steimle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejo,Z, Usmani,K, Valas,R, Vera,V, Villasana,D, Waldron,L, Walker,B, Wang,J, Wang,O, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,R, Wleczyk,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,S, Zhao,S, Dunn,D, von Niederhausern,A, Weiss,R, Smith,D, R, Holt,R,A, Smith,H,O,

TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
Weinstock,G, and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 478059)
Morley,K,C.
Direct Submission
Submitted (05-PEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (11-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Oct 9, 2002 this sequence version replaced gi:22795157.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GPH2
Center clone name: CH230-255D23

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 380245 bases at least Q40
Consensus quality: 385911 bases at least Q30
Consensus quality: 389757 bases at least Q20
Estimated insert size: 390792; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
NOTE: This sequence may represent more than one clone.
NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence, as soon as it is available and the accession number will be preserved.

Location/Qualifiers	1	235023	contig of 235023 bp in length
235024	235123	gap of unknown length	
235124	246065	contig of 10942 bp in length	
246066	246165	gap of unknown length	
246166	474601	contig of 228436 bp in length	
474602	474701	gap of unknown length	
474702	475761	contig of 1060 bp in length	
475762	475861	gap of unknown length	
475862	478059	contig of 2198 bp in length	

FEATURES

source

1. 478059

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-255D23"

misc_feature

1. 2769
/note="wgs_end_extension
clone_end:5p6"

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 259772)

Muzny,D,Marie,, Metzker,M, Lee,, Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benhamed,F., Biswalto,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Bunay,C., Butch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroli,L., De Anda,C., Dederich,D., Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Diya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagge,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gunardacne,F., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Haylak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M., Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kows,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Loshenshew,L., Louised,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mahoney,S., McLeod,M.P., McNeill,T.Z., Meenan,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Muntadas,M., Murphy,M., Naif,L., Nankervis,C., Neal,D., Newton,N., Nuyasa,N., Norris,S., Nwankweliem,O., Okunolu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Plankhock,C., Plopper,F., Poindexter,A., Popovic,D., Primm,E., Pu,L., Puato,M., Quito,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,M., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,D., Steimle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlarczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE
JOURNAL

Unpublished

2 (bases 1 to 259772)

REFERENCE
AUTHORS

Worley,K.C.

TITLE
JOURNAL

Submitted (14-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 259772)

Rat Genome Sequencing Consortium.

REFERENCE
AUTHORS

Submitted (08-OCT-2002)

TITLE
JOURNAL

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Sep 14, 2002 this sequence version replaced gi:21727094. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GONR

Center clone name: CH230-36G6

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 23717 bases at least Q40

Consensus quality: 238622 bases at least Q30

Estimated insert size: 252037; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

- * NOTE: Estimated insert size may differ from sequence length
- * (see http://www.hgsc.bcm.tmc.edu/docs/Gendank_draft_data.html)
- * NOTE: This sequence may represent more than one clone.
- * NOTE: This is a 'working draft' sequence. It currently
- * consists of 1 contig. Gaps between the contigs
- * are represented as runs of N. The order of the pieces
- * is believed to be correct as given, however the sizes
- * of the gaps between them are based on estimates that have
- * provided by the subcloner.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

1 259772: contig of 259772 bp in length.

Location/Qualifiers

1..259772

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-36G6"

misc_feature

1..1403

/note="wgs_contig"

misc_feature

3446..6094

/note="wgs_contig"

misc_feature

254751..255855

/note="wgs_contig"

misc_feature

258597..259772

/note="wgs_contig"

BASE COUNT

61631 a 50763 c 55648 g 72338 t 19392 others

ORIGIN

US-10-081-935-4 (1-194) x AC106943 (1-259772)

Alignment Scores:

Pred. No.: 56

Score: 159.50

Percent Similarity: 52.86%

Best Local Similarity: 51.43%

Query Match: 15.59%

DB: 2

Gaps: 1

US-10-081-935-4 (1-194) x AC106943 (1-259772)

84 SerTybAlaGlyValProileapleapleProProValaapThThAlaAlaProgluPro 103

Db 47749 AGCAAGCCACGACCGCCTCAGCGCAGCC-----ACGACACCGCCACGCA 47699

Qy 104 SerTybAlaGlyValProileapleapleProProValaapThThAlaAlaProgluPro 123

Db 47698 CCGCCACGACCGCCACGACCGCAGCGCAGCGCAGCGCAGCGCAGCGCAGCGCA 47639

Qy 124 GluProThrAlaGluProThrThiagluProThrgluGluProThrAlaValProThrgly 143

ACCESSION	AR178096	GI:17920989
VERSION	AR178096.1	
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	1 (bases 1 to 9636)	
AUTHORS	Williamson, K.C. and Kaslow, D.C.	
TITLE	Cloning and expression of plasmidum falciparum transmembrane-blocking target antigen, Pf5230	
JOURNAL	Patent: US 6316000-A 13-NOV-2001;	
FEATURES	Location/Qualifiers	
source	1..9636	
BASE COUNT	4156 a 974 c 1422 g 3084 t	
ORIGIN	/organism="unknown"	
Alignment Scores:		
Pred. No.:	3.26	Length: 9636
Score:	158.00	Matches: 35
Percent Similarity:	46.02%	Conservative: 17
Best Local Similarity:	30.97%	Mismatches: 31
Query Match:	15.44%	Indels: 30
DB:	6	Gaps: 3
US-10-081-935-4 (1-194) x AR178096 (1-9636)		
Qy	57 GlyGlnIleThrProCysValGluLysIleCysProLeuAspAlaArgIleSerValSer	76
Db	1515 GGATATATTTCACCTTGC-----CTTCTTTTTC	1486
Qy	77 AsnIleValIleAspGlnCysSerIysAlaGlyValProIleAspIleProValAsp	96
Db	1485 TCTACATAT-----TCACCTCTCTTTCACCTTACCTTACCTCTTCTTCTTCA	1438
Qy	97 ThrThrAlaAlaProGluProSerGluThrAlaGluProThrAlaGluProThrGluGlu	116
Db	1437 CCTACCTCTTACACCTTCCCTTACCTTCTTCTTCTTCTTCTTCTTCA	1378
Qy	117 ProThrAlaGluProThrAlaGluProThrAlaGluProThrAlaGluProThrGluGlu	136
Db	1377 CTTACCTCTTACCTTACCTTCTTACCTTCTTACCTTCTTACCTTCTTCA	1318
Qy	137 ProThrAlaValProThrGlyThrGlyGlyGlyValProThrGlyThrGlySerPheThr	156
Db	1317 CCTACCTCTTACCTTCTTCA-----	1294
Qy	157 ValThrGlyArgProThrAlaSerThrProAlaGluPhe	169
Db	1293 -----CCTACCTCTTACCTTCAATCTTCATTT	1267
RESULT 38		
LOCUS	195887	9636 bp DNA linear PAT 01-DEC-1998
DEFINITION	Sequence 1 from patent US 5735772.	
ACCESSION	195887	
VERSION	195887.1	GI:3940357
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	1 (bases 1 to 9636)	
AUTHORS	Williamson, K.C. and Kaslow, D.C.	
TITLE	Cloning and expression of plasmidum falciparum transmembrane blocking target antigen, Pf5230	
JOURNAL	Patent: US 5735772-A 13-MAR-1998;	
FEATURES	Location/Qualifiers	
source	1..9636	
BASE COUNT	4156 a 974 c 1422 g 3084 t	
ORIGIN	/organism="unknown"	
Alignment Scores:		

Pred. No.:	3.26	Length:	9636
Score:	158.00	Matches:	35
Percent Similarity:	46.02%	Conservative:	17
Best Local Similarity:	30.97%	Mismatches:	31
Query Match:	15.44%	Indels:	30
DB:	6	Gaps:	3
US-10-081-935-4 (1-194) x 195887 (1-9636)			
QY	57	GLYTHLLETHRRPCQGVVALGLUUALACYSPPROLEUABPILAAVGLISERVALSER	76
DB	1515	GGAAATATATTACCTTGC-----	1486
QY	77	ASNLLEVALAAGPGLNCYSESELYVALAGLYVALPROLLEAAPLIEPROVALASP	96
DB	1405	TCTACATAT-----TCACTCTCTCTTCCACTTCACTCTTCACTCTTCTTCA	1433
QY	97	ThrThraAlaAlaProGluProSerGluThraAlaGluProThraAlaGluProThrGluGlu	116
DB	1437	CCTACCTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCTTCTTCACTTCACTTCA	1376
QY	117	ProthraAlaGluProthraAlaGluProthraAlaGluProthraAlaGluProthrGluGlu	136
DB	1377	CTTACCTTCTTCACTTCACTTCTTCACTTCACTTCACTTCTTCACTTCACTTCACTTCA	1318
QY	137	ProthraAlaValProthrglyThrglyGlyGlyValProthrglyThrglySerPheThr	156
DB	1317	CCTACCTCTTCACTTCACTTCA-----	1299
QY	157	ValthrglyArgProthraAlaSerThrProAlaGluPhe	169
DB	1293	-----CCTACCTCTTCACTTCACTTCACTTCACTT	1267
RESULT 39			
LOCUS	PFAS230A	9654 bp	mRNA linear INV 29-DEC-1994
DEFINITION	Plasmodium falciparum (clone sy) 8230 mRNA, complete cds.		
ACCESSION	L04162		
VERSION	L04162.1	GI:605630	
KEYWORDS			
ORGANISM	Plasmodium falciparum 3D7		
SOURCE	Plasmodium falciparum 3D7		
REFERENCE	1 (bases 1 to 9654)		
AUTHORS	Bhatti,S., Alano,P., Luo,C., Hansra,S., Aikawa,M., Carter,R. and Elliott,J.F.		
TITLE	Gene cloning of a large Plasmodium falciparum sexual stage surface antigen		
JOURNAL	Unpublished (1993)		
COMMENT	Original source text: Plasmodium falciparum (strain 3D7) sexual stage and gametocyte possibly in asexual stages cDNA to mRNA.		
FEATURES	Location/Qualifiers		
source	1..9654		
	/organism="Plasmodium falciparum 3D7"		
	/mol_type="mRNA"		
	/strain="3D7"		
	/db_xref="taxon:36329"		
	/map="Chromosome 2"		
	/clone="sy"		
	/dev_stage="sexual stage and gametocyte possibly in asexual stages"		
	/germline		
gene	1..9654		
	/gene="8230"		
mRNA	1..9654		
	/gene="8230"		
5' UTR	1..179		
	/gene="8230"		
	179..9586		
	/gene="8230"		
CDS	/codon_start=1		
	/protein_id="AAAS7559.1"		
	/db_xref="GI:605631"		

QY 57 GYGLINILETHRPROCYSAVALGIUCLUALACYSEROLEUASPAALARGLESERVALSER 76
DB 1545 GGATATATATTCACCTTGC-----CTTCTTTTCA 1516

```
repeat_region
    /rpt_type=candem
    /rpt_unit=(A)n
    <1..44
repeat_region
    /rpt_type=candem
```


repeat_region 10502..10574
 /rpt_type=tandem
 /rpt_unit=(TAAA)n
 repeat_region 10502..10546

Alignment Scores:

Pred. No.:	4.96	Length:	15148
Score:	158.00	Matches:	35
Percent Similarity:	46.02%	Conservative:	17
Best Local Similarity:	30.97%	Mismatches:	31
Query Match:	15.44%	Indels:	30
DB:	3	Gaps:	3

US-10-081-935-4 (1-194) x AE001393 (1-15148)

QY	57	GLYGLNLEThrProCysValGluGluAlaCysProLeuAspAlaArgIleSerValSer	76
		:::	:::
DB	2236	GGATATATTTACCTTC	2207
		-----CTTCTTTTCA	
QY	77	AsnIleValValaspGlnCysSerLysAlaGlyValProIleaspIleProProValAsp	96
		:::	
DB	2206	TCTACATAT-----TCACCTCTCTCTTCACTTCACTTCACTTCTTCTTCA	2159
QY	97	ThrThraIaAlaProGluProSerGluThraIaGluProThraIaGluProThrgluGlu	116
		:::	:::
DB	2158	CCTACCTCTTCACTTCACTTCACTTCACTTCACTTCTTCTTCACTTCACTTCTTCA	2099
QY	117	ProThraIaGluProThraIaGluProThraIaGluProThrhIsgluProThrgluGlu	136
		:::	:::
DB	2098	CCTACCTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA	2039
QY	137	ProThraIaValProThrglyThrglyGlyGlyValProThrglyThrglySerPheThr	156
		:::	:::
DB	2038	CCTACCTCTTCACTTCACTTCTTCA-----	2015
QY	157	ValThrglyArgProThraIaSerThrProAlaGluPhe	169
		:::	:::
DB	2014	-----CCTACCTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT	1988

Search completed: December 6, 2003, 03:09:16
 Job time : 2869 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 13, 2003, 11:51:07 / Search time 52 Seconds

(without alignments)
152.786 Million cell updates/sec

Title: US-10-081-935-2

Perfect score: 82

Sequence: 1 M0FSHALITALVAAGLASA 18

Scoring table:

BLOSUM62		
Xgapop 10.0	Xgapext 0.5	
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10081935/runat_12112003_144407_3656/app_query.fasta_1.199
-DB=Issued_Patents_NA -QFMT=faststep -SUFFIX=rm1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi
-List=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10081935 @CGN 1.1 44 @runat_12112003_144407_3656 -NCPU=6 -ICPU=3
-NO MMAP -LARGEMEMORY -NEG SCORES=0 -WAIT -DSFLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6C.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	57.3	4403765	3 US-09-103-840A-2	Sequence 2, Appli
2	47	57.3	4411529	3 US-09-103-840A-1	Sequence 1, Appli
3	44	53.7	248	1 US-08-525-697-3	Sequence 3, Appli
4	44	53.7	1302	1 US-08-525-697-1	Sequence 1, Appli
5	44	53.7	10091	3 US-09-058-489-34	Sequence 34, Appli
6	42	51.2	471	4 US-09-252-991A-8612	Sequence 8612, Ap
7	42	51.2	504	4 US-09-134-001C-959	Sequence 959, App
8	42	51.2	1781	2 US-08-389-564B-21	Sequence 21, Appl
9	42	51.2	1781	2 US-08-389-564B-21	Sequence 21, Appl
10	42	51.2	1781	2 US-08-466-047B-22	Sequence 22, Appl
11	42	51.2	1781	3 US-08-466-047B-22	Sequence 22, Appl
12	42	51.2	2218	2 US-08-389-564B-18	Sequence 18, Appl

C 13	42	51.2	2218	3 US-09-066-597-1	Sequence 1, Appli
C 14	42	51.2	2218	3 US-08-466-047B-18	Sequence 18, Appl
C 15	42	51.2	2266	2 US-08-389-564B-20	Sequence 20, Appl
C 16	42	51.2	2266	3 US-08-466-047B-20	Sequence 20, Appl
C 17	42	51.2	4176	3 US-09-254-733-1	Sequence 1, Appli
C 18	42	51.2	8075	3 US-08-374-077C-1	Sequence 1, Appli
C 19	42	51.2	8075	3 US-08-895-590-1	Sequence 1, Appli
C 20	42	51.2	8075	4 US-05-539-879A-1	Sequence 1, Appli
C 21	42	51.2	8096	3 US-09-058-489-33	Sequence 3, Appli
C 22	41	50.0	528	4 US-09-252-991A-9137	Sequence 9137, Ap
C 23	41	50.0	972	4 US-09-252-991A-15702	Sequence 15702, A
C 24	41	50.0	975	4 US-09-252-991A-15674	Sequence 15674, A
C 25	41	50.0	1056	4 US-09-252-991A-8728	Sequence 8728, Ap
C 26	41	50.0	1218	4 US-09-552-322-3	Sequence 3, Appli
C 27	41	50.0	1230	4 US-09-328-352-1918	Sequence 1918, Ap
C 28	41	50.0	1485	4 US-09-252-991A-15656	Sequence 15656, A
C 29	41	50.0	1566	4 US-09-252-991A-9220	Sequence 9220, Ap
C 30	41	50.0	1710	4 US-09-252-991A-9057	Sequence 9057, Ap
C 31	41	50.0	2123	4 US-09-724-864-18	Sequence 18, Appl
C 32	41	50.0	5045	3 US-09-390-721-1	Sequence 1, Appli
C 33	41	50.0	5045	3 US-09-390-721-3	Sequence 3, Appli
C 34	41	50.0	36412	4 US-08-311-731A-132	Sequence 132, Appl
C 35	40	48.8	304	4 US-09-313-294A-2891	Sequence 2891, Ap
C 36	40	48.8	465	4 US-09-252-991A-16275	Sequence 16275, A
C 37	40	48.8	504	4 US-09-512-342-19	Sequence 19, Appl
C 38	40	48.8	669	4 US-09-252-991A-3703	Sequence 3703, Ap
C 39	40	48.8	717	3 US-08-913-014A-7	Sequence 7, Appli
C 40	40	48.8	765	4 US-09-252-991A-9931	Sequence 9931, Ap
C 41	40	48.8	795	3 US-08-913-014A-8	Sequence 8, Appli
C 42	40	48.8	951	4 US-09-252-991A-16386	Sequence 16386, A
C 43	40	48.8	1053	4 US-09-252-991A-6421	Sequence 6421, Ap
C 44	40	48.8	1092	4 US-09-252-991A-16212	Sequence 16212, A
C 45	40	48.8	1179	4 US-09-252-991A-10081	Sequence 10081, A

ALIGNMENTS

```
RESULT 1
US-09-103-840A-2
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: FRASER, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Alignment Scores:
Pred. No.: 1.85e+05
Score: 47.00
Percent Similarity: 92.86%
Best Local Similarity: 64.29%
Query Match: 57.32%
DB: 3
Length: 4403765
Matches: 9
Conservative: 4
Mismatches: 1
Indels: 0
Gaps: 0
US-10-081-935-2 (1-18) x US-09-103-840A-2 (1-4403765)
```

Oy 3 PhseSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAla 16
 Db 1218600 TTCCGCGACCCCTGCTGTCGATCGTTGCCCGCTGATTTGGC 1218641

RESULT 2
 US-09-103-840A-1
 Sequence 1, Application US/09103840A
 Patent No. 6294328
 GENERAL INFORMATION:
 APPLICANT: FLEISCHMAN, Robert D.
 APPLICANT: WHITE, Owen R.
 APPLICANT: FRASER, Claire M.
 APPLICANT: VENTER, John C.
 TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 TITLE OF INVENTION: TUBERCULOSIS
 FILE REFERENCE: 24366-20007.00
 CURRENT APPLICATION NUMBER: US/09/103,840A
 CURRENT FILING DATE: 1998-06-24
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
 LENGTH: 4411529
 TYPE: DNA
 ORGANISM: Mycobacterium tuberculosis
 OTHER INFORMATION: H37Rv
 US-09-103-840A-1

Alignment Scores:
 Pred. No.: 1 85e+05 length: 4411529
 Score: 47.00 Matches: 9
 Percent Similarity: 92.86% Conservative: 4
 Best Local Similarity: 64.29% Mismatches: 1
 Query Match: 57.32% Indels: 0
 DB: 3 Gaps: 0

US-10-081-935-2 (1-18) x US-09-103-840A-1 (1-4411529)
 Oy 3 PhseSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAla 16
 Db 1219072 TTCCGCGACCCCTGCTGTCGATCGTTGCCCGCTGATTTGGC 1219113

RESULT 3
 US-08-525-697-3
 Sequence 3, Application US/08525697
 Patent No. 5795764
 GENERAL INFORMATION:
 APPLICANT: Christgau, Stephan
 APPLICANT: Andersen, Lene N
 APPLICANT: Kauppinen, Sakari
 APPLICANT: Heldt-Hansen, Hans P
 APPLICANT: Dalboege, Henrik
 TITLE OF INVENTION: AN ENZYME EXHIBITING MANNANASE ACTIVITY
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 57957640 No. 5795764disk of No. 5795764th America, Inc
 STREET: 405 Lexington Avenue, 64th Floor
 CITY: New York
 STATE: New York
 COUNTRY: United States of America
 ZIP: 10174-6401
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/525,697
 FILING DATE: 21-SEP-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Harrington, James J.
 REGISTRATION NUMBER: 38,711
 REFERENCE/DOCKET NUMBER: 4004.204-US

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DK 0486/93
: FILING DATE: 30-APR-1993
: CLASSIFICATION: 435
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 248 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-525-697-3

Alignment Scores:
Pred. No.: 3.04 Length: 248
Score: 44.00 Matches: 8
Percent Similarity: 77.78% Conservative: 6
Best Local Similarity: 44.44% Mismatches: 4
Query Match: 53.66% Indels: 0
DB: 1 Gaps: 0

US-10-081-935-2 (1-18) x US-08-525-697-3 (1-248)

Cy 1 MetGlnPhseSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
Db 24 ATGAGCTTTCTACATGCTCTCTCAGACCTCCGACACCTGGGGGATGGCAGCGCT 77

RESULT 4
US-08-525-697-1
: Sequence 1, Application US/08525697
: Patent No. 5795764
: GENERAL INFORMATION:
: APPLICANT: Christgau, Stephan
: APPLICANT: Andersen, Lene N
: APPLICANT: Kauppinen, Sakari
: APPLICANT: Heidt-Hansen, Hans P
: APPLICANT: Dalboege, Henrik
: TITLE OF INVENTION: AN ENZYME EXHIBITING MANNANASE ACTIVITY
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NO. 57957640 NO. 5795764disk of No. 5795764th America, Inc
: STREET: 405 Lexington Avenue, 64th Floor
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/525,697
: FILING DATE: 21-SEP-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Harrington, James J.
: REGISTRATION NUMBER: 38,711
: REFERENCE/DOCKET NUMBER: 4004.204-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DK 0486/93
: FILING DATE: 30-APR-1993
: CLASSIFICATION: 435
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1302 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 61..1192
US-08-525-697-1

Alignment Scores:
Pred. No.: 25
Score: 44.00
Percent Similarity: 77.78%
Best Local Similarity: 44.44%
Query Match: 53.66%
DB: 1
Gaps: 0

US-10-081-935-2 (1-18) x US-08-525-697-1 (1-1302)

Qy 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
Db 61 ATGAAGCTTCTTCATGCTCTCCACGCTCGCCAGCCTGGGGGTGGCGAGCGCT 114

RESULT 5
US-09-058-489-34
; Sequence 34, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WH197-08PA
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 10091
; TYPE: DNA
; ORGANISM: Human
US-09-058-489-34

Alignment Scores:
Pred. No.: 339
Score: 44.00
Percent Similarity: 76.47%
Best Local Similarity: 52.94%
Query Match: 53.66%
DB: 3
Gaps: 0

US-10-081-935-2 (1-18) x US-09-058-489-34 (1-10091)

Qy 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSer 17
Db 3483 CTTCAACAAATCATGCTTGTACTTGTGACGAGAAACCTTGCAACC 3533

RESULT 6
US-09-252-991A-8612
; Sequence 8612, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
```

```

; SEQ ID NO 8612
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8612

Alignment Scores:
Pred. No.: 168
Score: 42.00
Percent Similarity: 76.47%
Best Local Similarity: 41.18%
Query Match: 51.22%
DB: 4
Gaps: 0

US-10-081-935-2 (1-18) x US-09-252-991A-8612 (1-471)

Qy 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSer 17
Db 70 ATGAATTCACGATTCATCGCGGCATCGTTGGCCGCCACTGTCACAC 120

RESULT 7
US-09-134-001C-959/C
; Sequence 959, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucetre-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 959
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-959

Alignment Scores:
Pred. No.: 183
Score: 42.00
Percent Similarity: 82.35%
Best Local Similarity: 58.82%
Query Match: 51.22%
DB: 4
Gaps: 1

US-10-081-935-2 (1-18) x US-09-134-001C-959 (1-504)

Qy 2 GlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
Db 134 CAATTTCACATTCCTTCTACGCTTAATCGA-----TTGGCCATCATCT 90

RESULT 8
US-08-389-564B-21/C
; Sequence 21, Application US/08389564B
; Patent No. 5989870
; GENERAL INFORMATION:
; APPLICANT: Nakari, Tiina H.
; APPLICANT: Onnela, Maija-Leena
; APPLICANT: Iim n, Marja H.
; APPLICANT: Penttil, Merja E.
; TITLE OF INVENTION: A METHOD FOR CLONING ACTIVE PROMOTERS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
```

```

;
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/389,564B
; FILING DATE: 16-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/932,485
; FILING DATE: 19-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 86 10600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, GRANT E.
; REGISTRATION NUMBER: 41,264
; REFERENCE/DOCKET NUMBER: 1716.008000G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1781 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
;
; US-08-389-564B-21
;
; Alignment Scores:
; Pred. No.: 91.3 length: 1781
; Score: 42.00 Matches: 7
; Percent Similarity: 84.62% Conservative: 4
; Best Local Similarity: 53.85% Mismatches: 2
; Query Match: 51.22% Indels: 0
; DB: 2 Gaps: 0
;
; US-10-081-935-2 (1-18) x US-08-389-564B-21 (1-1781)
;
; QY 3 pHeSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeu 15
; DB 1031 TTAGCCACTTGATGTGTACATTATTAGCCGCTGTATA 993
;
; RESULT 9
; US-08-389-564B-22/C
; Sequence 22, Application US/08389564B
; Patent No. 5989870
; GENERAL INFORMATION:
; APPLICANT: Nakari, Tiina H.
; APPLICANT: Onnela, Maija-Leena
; APPLICANT: Ilm n, Marja H.
; APPLICANT: Penttil, Merja E.
; TITLE OF INVENTION: A METHOD FOR CLONING ACTIVE PROMOTERS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; SOFTWARE: Patentin Release #1.0, Version #1.25
; COMPUTER: IBM PC compatible
```

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/389,564B
; FILING DATE: 16-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/932,485
; FILING DATE: 19-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 86 10600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, GRANT E.
; REGISTRATION NUMBER: 41,264
; REFERENCE/DOCKET NUMBER: 1716.008000G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1781 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
;
; US-08-389-564B-22
;
; Alignment Scores:
; Pred. No.: 91.3 length: 1781
; Score: 42.00 Matches: 7
; Percent Similarity: 84.62% Conservative: 4
; Best Local Similarity: 53.85% Mismatches: 2
; Query Match: 51.22% Indels: 0
; DB: 2 Gaps: 0
;
; US-10-081-935-2 (1-18) x US-08-389-564B-22 (1-1781)
;
; QY 3 pHeSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeu 15
; DB 1031 TTAGCCACTTGATGTGTACATTATTAGCCGCTGTATA 993
;
; RESULT 10
; US-08-466-047B-21/C
; Sequence 21, Application US/08466047B
; Patent No. 6011147
; GENERAL INFORMATION:
; APPLICANT: Nakari, Tiina H.
; APPLICANT: Onnela, Maija-Leena
; APPLICANT: Ilm n, Marja H.
; APPLICANT: Nevalainen, Kaisa Milja Helena
; APPLICANT: Penttil, Merja E.
; TITLE OF INVENTION: Fungal Promoters Active In The Presence
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```



```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/389,564B
: FILING DATE: 16-FEB-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/932,485
: FILING DATE: 19-AUG-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/496,155
: FILING DATE: 19-MAR-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/044,077
: FILING DATE: 29-APR-1987
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 86 10600
: FILING DATE: 30-APR-1986
: ATTORNEY/AGENT INFORMATION:
: NAME: REED, GRANT E.
: REGISTRATION NUMBER: 41,264
: REFERENCE/DOCKET NUMBER: 1716.008000G
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 18:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: both
: MOLECULE TYPE: CDNA
:
: US-08-389-564B-18
:
: Alignment Scores:
: Pred. No.: 121 Length: 2218
: Score: 42.00 Matches: 7
: Percent Similarity: 84.62% Conservative: 4
: Best Local Similarity: 53.85% Mismatches: 2
: Query Match: 51.22% Indels: 0
: DB: 2 Gaps: 0
:
: US-10-081-935-2 (1-18) x US-08-389-564B-18 (1-2218)
:
: QY 3 PheseRHisAlaLeuIleAlaLeuValAlaAglIleu 15
: Db 1466 TTAGCCACTTGATTGTACATATTATAGCCGCTGTATA 1428
:
: RESULT 13
: US-09-066-597-1/c
: Sequence 1, Application US/09066597
: Patent No. 6001595
: GENERAL INFORMATION:
: APPLICANT: IIm n, Marija
: APPLICANT: Onnela, Marija-Leena
: APPLICANT: Penttil, Merja
: TITLE OF INVENTION: New Promoters and Uses Thereof
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
: STREET: 1100 New York Avenue, NW, Suite 600
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20005-3934
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/066,597
: FILING DATE: Herewith
: CLASSIFICATION:
: PRIOR APPLICATION DATA:

```

```

: APPLICATION NUMBER: US 60/032,156
: FILING DATE: 29-NOV-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/032,959
: FILING DATE: 13-DEC-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/040,140
: FILING DATE: 10-MAR-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/980,061
: FILING DATE: 26-NOV-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/FI97/00742
: FILING DATE: 01-DEC-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Kim, Judith U.
: REGISTRATION NUMBER: 40,679
: REFERENCE/DOCKET NUMBER: 1716.0660004/MAC/JUK
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
:
: US-09-066-597-1
:
: Alignment Scores:
: Pred. No.: 121 Length: 2218
: Score: 42.00 Matches: 7
: Percent Similarity: 84.62% Conservative: 4
: Best Local Similarity: 53.85% Mismatches: 2
: Query Match: 51.22% Indels: 0
: DB: 3 Gaps: 0
:
: US-10-081-935-2 (1-18) x US-09-066-597-1 (1-2218)
:
: QY 3 PheseRHisAlaLeuIleAlaLeuValAlaAglIleu 15
: Db 1466 TTAGCCACTTGATTGTACATATTATAGCCGCTGTATA 1428
:
: RESULT 14
: US-08-466-047B-18/c
: Sequence 18, Application US/08466047B
: Patent No. 6011147
: GENERAL INFORMATION:
: APPLICANT: Nakari, Tiina H.
: APPLICANT: Onnela, Marija-Leena
: APPLICANT: IIm n, Marija H.
: APPLICANT: Nevalainen, Kaisu
: APPLICANT: Penttil, Merja E.
: TITLE OF INVENTION: Fungal Promoters Active In The Presence
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
: STREET: 1100 New York Avenue, Suite 600
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/466,047B
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/389,564
; FILING DATE: 16-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/932,564
; FILING DATE: 19-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 86 10600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, GRANT E.
; REGISTRATION NUMBER: 41,264
; REFERENCE/DOCKET NUMBER: 1716.008000H
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: CDNA
; US-08-466-047B-18

Alignment Scores:
Pred. No.: 121 Length: 2218
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 51.22% Indels: 0
Gaps: 0

US-10-081-935-2 (1-18) x US-08-466-047B-18 (1-2218)

QY 3 PheSerHisAlaLeuIleAlaLeuValAlaIleGlyLeu 15
Db 1466 TTTACCACTTGATGTACAAATTATTAGCCGCTGGTATA 1428

RESULT 15
US-08-389-564B-20/C
; Sequence 20, Application US/08389564B
; GENERAL INFORMATION:
; PATENT NO. 5989870
; APPLICANT: Nakari, Tiina H.
; APPLICANT: Onnela, Maija-Leena
; APPLICANT: Ilm n, Marja H.
; APPLICANT: Penttil, Merja E.
; TITLE OF INVENTION: A METHOD FOR CLONING ACTIVE PROMOTERS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/389,564B
; FILING DATE: 16-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

```

; APPLICATION NUMBER: US 07/932,485
; FILING DATE: 19-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 86 10600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, GRANT E.
; REGISTRATION NUMBER: 41,264
; REFERENCE/DOCKET NUMBER: 1716.008000G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2266 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: CDNA
; US-08-389-564B-20

Alignment Scores:
Pred. No.: 124 Length: 2266
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 51.22% Indels: 0
Gaps: 0

US-10-081-935-2 (1-18) x US-08-389-564B-20 (1-2266)

QY 3 PheSerHisAlaLeuIleAlaLeuValAlaIleGlyLeu 15
Db 1516 TTTACCACTTGATGTACAAATTATTAGCCGCTGGTATA 1478

RESULT 16
US-08-466-047B-20/C
; Sequence 20, Application US/08466047B
; GENERAL INFORMATION:
; PATENT NO. 601147
; APPLICANT: Nakari, Tiina H.
; APPLICANT: Onnela, Maija-Leena
; APPLICANT: Ilm n, Marja H.
; APPLICANT: Nevalainen, Kaisa Milja Helena
; APPLICANT: Penttil, Merja E.
; TITLE OF INVENTION: Fungal Promoters Active In The Presence
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,047B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/389,564
; FILING DATE: 16-FEB-1995
```



```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/932,564
; FILING DATE: 19-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 86 10600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, GRANT E.
; REGISTRATION NUMBER: 41,264
; REFERENCE/DOCKET NUMBER: 1716.008000H
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2266 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; US-08-466-047B-20

Alignment Scores:
Pred. No.: 124 Length: 2266
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 51.22% Indels: 0
DB: 3 Gaps: 0

US-10-081-935-2 (1-18) x US-08-466-047B-20 (1-2266)
QY 3 PhseSerHisAlaLeuIleAlaValAlaAlaGlyLeu 15
Db 1516 TTTAGCCACTTAATTGTACATTATTAGCGCGTGTATA 1478

RESULT 17
US-09-254-733-1/c
; Sequence 1, Application US/09254733
; Patent No. 6277596
; GENERAL INFORMATION:
; APPLICANT: WATANABE, MANABU
; APPLICANT: MORIYA, TATSUKI
; APPLICANT: AOYAGI, KAORU
; APPLICANT: SUMIDA, NAOMI
; APPLICANT: MURAKAMI, TAKESHI
; TITLE OF INVENTION: REGULATORY SEQUENCE OF CELLULASE cbh1 GENES ORIGINATING
; TITLE OF INVENTION: IN TRICHODERMA VIRIDE AND SYSTEM FOR MASS-PRODUCING
; FILE REFERENCE: 99-0266*/LC(IMC)/00144
; CURRENT APPLICATION NUMBER: US/09/254,733
; FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4176
; TYPE: DNA
; ORGANISM: TRICHODERMA VIRIDE MC300-1
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (1438) .. (1488)
; NAME/KEY: CDS
; LOCATION: (1438) .. (1488)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1489) .. (3108)
```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1489) .. (1898)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (1899) .. (1965)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1966) .. (2662)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (2663) .. (2724)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2725) .. (3108)
; US-09-254-733-1

Alignment Scores:
Pred. No.: 270 Length: 4176
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 51.22% Indels: 0
DB: 3 Gaps: 0

US-10-081-935-2 (1-18) x US-09-254-733-1 (1-4176)
QY 3 PhseSerHisAlaLeuIleAlaValAlaAlaGlyLeu 15
Db 708 TTTAGCCACTTAATTGTACATTATTAGCGCGTGTATA 670

RESULT 18
US-08-374-077C-1/c
; Sequence 1, Application US/08374077C
; Patent No. 6027912
; GENERAL INFORMATION:
; APPLICANT: Hall, Linda M.
; APPLICANT: Ren, Dejian
; APPLICANT: Zheng, Wei
; APPLICANT: Dubald, Manuel Marcel Paul
; TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha1
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314-3187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,077C
; FILING DATE: 19-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm M.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 022650-264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
```

FEATURE:
NAME/KEY: CDS
LOCATION: 157..7704
US-08-374-077C-1

Alignment Scores:
Pred. No.: 624
Score: 42.00
Percent Similarity: 83.33%
Best Local Similarity: 44.44%
Query Match: 51.22%
DB: 3
Gaps: 0

US-10-081-935-2 (1-18) x US-08-374-077C-1 (1-8075)

Qy 1 MetGlnPheSerHisAlaLeuValAlaIaGlyLeuAlaSerAla 18
Db 1917 TTGGCGTTTCCAGCGCGTTGTCTTTGGCCAGCGCACCTGCACGACG 1864

RESULT 19
US-08-895-590-1/C
Sequence 1, Application US/08895590
Patent No. 6207410
GENERAL INFORMATION:
APPLICANT: Hall, Linda M.
APPLICANT: Ren, Dejian
APPLICANT: Zheng, Wei
APPLICANT: Dubald, Manuel Marcel Paul
TITLE OF INVENTION: Genes Encoding an Insect Calcium Channel
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895,590
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/374,888
FILING DATE: 19-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-263
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8075 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 157..7704
US-08-895-590-1

Alignment Scores:
Pred. No.: 624
Score: 42.00
Percent Similarity: 83.33%
Best Local Similarity: 44.44%
Query Match: 51.22%
DB: 3
Gaps: 0

Query Match: 51.22%
DB: 3
Gaps: 0

US-10-081-935-2 (1-18) x US-08-895-590-1 (1-8075)

Qy 1 MetGlnPheSerHisAlaLeuValAlaIaGlyLeuAlaSerAla 18
Db 1917 TTGGCGTTTCCAGCGCGTTGTCTTTGGCCAGCGCACCTGCACGACG 1864

RESULT 20
US-09-539-879A-1/C
Sequence 1, Application US/09539879A
Patent No. 6436627
GENERAL INFORMATION:
APPLICANT: Hall, Linda M.
APPLICANT: Ren, Dejian
APPLICANT: Zheng, Wei
APPLICANT: Dubald, Manuel Marcel Paul
TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/539,879A
FILING DATE: 31-Mar-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/111,865
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/374,077
FILING DATE: 19-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-264
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8075 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 157..7704
US-09-539-879A-1

Alignment Scores:
Pred. No.: 624
Score: 42.00
Percent Similarity: 83.33%
Best Local Similarity: 44.44%
Query Match: 51.22%
DB: 4
Gaps: 0

US-10-081-935-2 (1-18) x US-09-539-879A-1 (1-8075)

Qy 1 MetGlnPheSerHisAlaLeuValAlaIaGlyLeuAlaSerAla 18

Db
1917 TTGGGTTTCCACGGCGTTGTCTTTGGCCACGGGCACCTGCAGCAGC 1864

```

RESULT 2)
US-09-056-489-33
; Sequence 33, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Lahn, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; TITLE OF INVENTION: the Y Chromosome
; FILE REFERENCE: WH197-08PA
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 8096
; TYPE: DNA
; ORGANISM: Human
US-09-058-489-33

```

Alignment Scores:	
Pred. No.:	626
Score:	42.00
Percent Similarity:	70.59%
Best Local Similarity:	47.06%
Query Match:	51.22%
DB:	3
Length:	8036
Matches:	8
Conservative:	4
Mismatches:	5
Indels:	0
Gaps:	0

US-10-081-935-2 (1-18) X US-09-058-489-33 (1-8096)

OY 1 MetGlnPheSerHisAlaLeuIleValAlaAlaGlyLeuAlaSer 17
 :::||| ::||| ||| |
Db 1854 CTTCAACACAATCATGCCCTAGTTCATTGGTAGCAGAATAACTTTAACT 1904

RESULT 22
 US-09-252-991A-9137/c
 : Sequence 9137, Application US/09252991A
 : Patent No. 6551795
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Marc J. Rubenfield et al.
 :
 : TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 : TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 : FILE REFERENCE: 107196.136
 :
 : CURRENT APPLICATION NUMBER: US/09/252,991A
 :
 : CURRENT FILING DATE: 1999-02-18
 :
 : PRIOR APPLICATION NUMBER: US 60/074,788
 :
 : PRIOR FILING DATE: 1998-02-18
 :
 : PRIOR APPLICATION NUMBER: US 60/094,190
 :
 : PRIOR FILING DATE: 1998-07-27
 :
 : NUMBER OF SEQ ID NOS: 33142
 :
 : SEQ ID NO 9137
 :
 : LENGTH: 528
 :
 : TYPE: DNA
 :
 : ORGANISM: *Pseudomonas aeruginosa*
 :
 : US-09-252-991A-9137

Alignment Scores:	
Pred. No.:	30.4
Score:	41.00
Percent Similarity:	61.1%
Best Local Similarity:	44.44%
Query Match:	50.00%
DB:	4
Length:	52
Matches:	8
Conservative:	3
Mismatches:	7
Indels:	0
Gaps:	0

US-10-081-935-2 (1-18) X US-09-252-991A-9137 (1-528)

QY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18

Db 437 CTGCAGTTGACACCACTACTTGGCGGAGTCGGCGACGCG 384

```

US-09-252-991A-15702
RESULT 23
; Sequence 15702, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15702
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (669)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-15702

```

Alignment Scores:	
Pred. No.:	66.1
Score:	41.00
Percent Similarity:	72.22%
Best Local Similarity:	61.11%
Query Match:	50.00%
DB:	4
Length:	979
Matches:	11
Conservative:	2
Mismatches:	5
Indels:	0
Gaps:	0

US-10-081-935-2 (1-18) x US-09-252-991A-15702 (1-972)

QY 1 MeGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
 ||||| :: ||||| ||||| ||||| :: |||
 DB 151 ATGCAGGGCGCGCGCCCTGATCCGCGTCGTCCGCGCCGGGAGAGCATGGCG 204

```

RESULT 24
US-09-252-991A-15674
: Sequence 15674, Application US/09252991A
: Patent No. 6551795
:
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: PRIORITY FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIORITY FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIORITY FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 15674
: LENGTH: 975
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (571)
: OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-15674

```

Alignment Scores:	66.3	Length:	979
Pred. No.:	41.00	Matches:	11
Score:	72.22%	Conservative:	2

Alignment Scores:

Pred. No.: 113 Length: 1485
Score: 41.00 Matches: 11
Percent Similarity: 72.22% Conservative: 2
Best Local Similarity: 61.11% Mismatches: 5
Query Match: 50.00% Indels: 0
DB: 4 Gaps: 0

US-10-081-935-2 (1-18) x US-09-252-991A-15656 (1-1485)

QY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
DB 933 ATGCAGAGCGCGCGCGCTGATCGCGCTGCGCGCGGAGAGCATGCGG 880

RESULT 29

US-09-252-991A-9220/C
; Sequence 9220, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9220
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9220

Alignment Scores:

Pred. No.: 121 Length: 1566
Score: 41.00 Matches: 8
Percent Similarity: 61.11% Conservative: 3
Best Local Similarity: 44.44% Mismatches: 7
Query Match: 50.00% Indels: 0
DB: 4 Gaps: 0

US-10-081-935-2 (1-18) x US-09-252-991A-9220 (1-1566)

QY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
DB 979 CTGCAGCTTGAACACCACTGCTGCGGATCGCGCGCTCGGCAGCGG 926

RESULT 30

US-09-252-991A-9057/C
; Sequence 9057, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9057
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9057

Alignment Scores:

Pred. No.: 136 Length: 1710
Score: 41.00 Matches: 8
Percent Similarity: 61.11% Conservative: 3
Best Local Similarity: 44.44% Mismatches: 7
Query Match: 50.00% Indels: 0
DB: 4 Gaps: 0

US-10-081-935-2 (1-18) x US-09-252-991A-9057 (1-1710)

QY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
DB 1143 CTGCAGCTTGAACACCACTGCTGCGGATCGCGCGCTCGGCAGCGG 1090

RESULT 31

US-09-724-864-18
; Sequence 18, Application US/09724864
; Patent No. 6380362
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; FILE REFERENCE: 11000.1050U1
; CURRENT APPLICATION NUMBER: US/09/724,864
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 2123
; TYPE: DNA
; ORGANISM: Mouse
US-09-724-864-18

Alignment Scores:

Pred. No.: 179 Length: 2123
Score: 41.00 Matches: 8
Percent Similarity: 66.67% Conservative: 4
Best Local Similarity: 44.44% Mismatches: 6
Query Match: 50.00% Indels: 0
DB: 4 Gaps: 0

US-10-081-935-2 (1-18) x US-09-724-864-18 (1-2123)

QY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAla 18
DB 870 CTGCCTTGAAGCCACCACTGCTGCTGCTGCGCAATGCGTGTGTACAGCA 923

RESULT 32

US-09-390-721-1/C
; Sequence 1, Application US/09390721
; Patent No. 6197591
; GENERAL INFORMATION:
; APPLICANT: STUTZMAN-ENGWALL, KIM J.
; TITLE OF INVENTION: STREPTOMYCES AVERMITILIS REGULATORY GENES FOR INCREASED
; FILE REFERENCE: PC9944A
; CURRENT APPLICATION NUMBER: US/09/390,721
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: 60/100,134
; EARLIER FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 5045
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1112)..(2317)
; OTHER INFORMATION: aver1 ORF

```
US-09-390-721-1
Alignment Scores:
Pred. No.: 537
Score: 41.00
Percent Similarity: 70.59%
Best Local Similarity: 47.06%
Query Match: 50.00%
DB: 3
Gaps: 0

US-10-081-935-2 (1-18) x US-09-390-721-1 (1-5045)
QY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaIleAGlyLeuAlaSer 17
DB 4382 CTGCACACTGACACACCGCTTGCTGCACTTCGTGAGCGCGGCTGCGCC 4332

RESULT 33
US-09-390-721-3/c
; Sequence 3, Application US/09390721
; Patent No. 6197591
; GENERAL INFORMATION:
; APPLICANT: STUTZMAN-ENGWALL, KIM J.
; APPLICANT: PRICE, BRENDA S.
; TITLE OF INVENTION: STREPTOMYCES AVERMITILIS REGULATORY GENES FOR INCREASED
; FILE REFERENCE: PC9944A
; CURRENT APPLICATION NUMBER: US/09/390,721
; EARLIER APPLICATION NUMBER: 60/100,134
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 5045
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2314)..(3021)
; OTHER INFORMATION: aver2 ORF
US-09-390-721-3

Alignment Scores:
Pred. No.: 537
Score: 41.00
Percent Similarity: 70.59%
Best Local Similarity: 47.06%
Query Match: 50.00%
DB: 3
Gaps: 0

US-10-081-935-2 (1-18) x US-09-390-721-3 (1-5045)
QY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaIleAGlyLeuAlaSer 17
DB 4382 CTGCACACTGACACACCGCTTGCTGCACTTCGTGAGCGCGGCTGCGCC 4332

RESULT 34
US-08-311-731A-132
; Sequence 132, Application US/08311731A
; Patent No. 6583286
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-1
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
```

```
ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36412 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-132

Alignment Scores:
Pred. No.: 6,636+03
Score: 41.00
Percent Similarity: 85.71%
Best Local Similarity: 57.14%
Query Match: 50.00%
DB: 4
Gaps: 0

US-10-081-935-2 (1-18) x US-08-311-731A-132 (1-36412)
QY 4 SerHisAlaLeuIleAlaLeuValAlaIleAGlyLeuAlaSer 17
DB 8097 ACTCAGCGCGCTTCGCGCACTCATCGTGGCTGACTCG 8138

RESULT 35
US-09-313-294A-2891/c
; Sequence 2891, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 2891
; LENGTH: 304
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID NO. 6476212 700553448H1
; NAME/KEY: unsure
; LOCATION: 138, 178
; OTHER INFORMATION: a, c, c, g, or other
US-09-313-294A-2891

Alignment Scores:
Pred. No.: 23.6
Score: 40.00
Length: 304
Matches: 9
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2003, 01:54:48 ; Search time 1932 Seconds

(without alignments)
2440.512 Million cell updates/sec

Title: US-10-081-935-4
Perfect score: 1023
Sequence: 1 MCFSHALIALVAGLASAQL.....NVRASVGINAALLGLAAVL 194

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O=/cgm2_3/USFTO.spool/US10081935/runat_04122003_134830_10487/app-query.fasta_1.391
-DB=EST -QPM=fastap -SUFFIX=ref -MINMATCH=0.1 -LOOPC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10081935 @CCN 1.1 2135 @runat_04122003_134830_10487 -NCPU=6 -ICPU=3
-NO_MMP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MKRN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_estb1:*
2: em_estb2:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pln:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_phg:*
27: em_gse_vr1:*
28: gb_gse1:*

29: gb_gse2:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1010	98.7	737	10	BF251103
2	706	68.0	557	10	BF252723
3	701	68.5	531	10	BF251334
4	646	63.1	395	10	BF251817
5	608	59.4	401	10	BF252856
6	451	44.1	395	10	BF251279
7	412.5	40.3	602	13	BO767421
8	399.5	39.1	538	13	BO767128
9	382	37.3	419	13	BO761104
10	322	31.5	311	13	BO499065
11	219.5	21.5	667	12	BM077163
12	202.5	19.8	564	12	BM666652
13	193	18.9	553	12	BM666040
14	185	18.1	952	14	CD458295
15	184	18.0	731	13	CD458239
16	181.5	17.7	688	12	BM872350
17	176	17.2	701	12	BM872513
18	172.5	16.9	640	14	CD035585
19	168.5	16.5	612	12	B3383609
20	168	16.4	1038	29	CNS05RNM
21	167	16.3	567	12	BM666206
22	167	16.3	607	13	BU064230
23	167	16.3	623	13	BU066221
24	167	16.3	634	13	BU065834
25	167	16.3	657	13	BU064409
26	167	16.3	664	13	BU065050
27	167	16.3	666	13	BU065008
28	167	16.3	699	13	BU066113
29	167	16.3	718	13	BU064737
30	167	16.3	726	13	BU064923
31	167	16.3	841	14	CD458964
32	167	16.3	867	14	CD458020
33	166	16.2	606	12	BM668215
34	166	16.2	636	13	BU064561
35	166	16.2	822	14	CD460963
36	164.5	16.1	611	13	BU064419
37	164.5	16.1	611	13	BU065692
38	164	16.0	612	9	AU284313
39	164	16.0	643	13	BU061554
40	164	16.0	874	14	CD459301
41	163	15.9	626	13	BU066074
42	161	15.7	869	14	CD458008
43	160.5	15.6	556	12	BM872159
44	160	15.6	531	6	AU191003
45	160	15.6	725	14	CD044514

ALIGNMENTS

RESULT 1
LOCUS BF251103
DEFINITION BF251103 Coccidioides immitis spherule CDNA library Coccidioides immitis CDNA clone CIAAC67 5' sequence, mRNA sequence.
VERSION BF251103
ACCESSION BF251103
KEYWORDS BF251103.1 GI:16931246
SOURCE EST.
ORGANISM Coccidioides immitis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE 1 (bases 1 to 737)

AUTHORS Gardner, M.J. and Kirkland, T.
TITLE Generation of ESTs from *Coccidioides immitis* spherule cDNA library
JOURNAL Unpublished
COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org

FEATURES
source Location/Qualifiers
1..737

/organism="Coccidioides immitis"
/mol_type="mRNA"
/db_xref="taxon:5501"
/clone="CIAC67"
/dev_stage="spherule"
/lab_host="SOLR"
/clone_lib="Coccidioides immitis spherule cDNA library"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 136 a 271 c 168 g 162 t
ORIGIN

Alignment Scores:
Pred. No.: 9,2e-69 Length: 737
Score: 1010.00 Matches: 194
Percent Similarity: 99.49% Conservative: 0
Best Local Similarity: 99.49% Mismatches: 0
Query Match: 98.73% Indels: 1
DB: 10 Gaps: 0

US-10-081-935-4 (1-194) x BF251103 (1-737)

```
OY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu 20
DB 99 ATGCAgTTCCTCAGCGCTCTCATGCTCGTGGCGGCGCTCGCAGATCCAGGCTC 158
OY 21 ProAapIleProProCysAlaLeuAsnCysPheValGluAlaLeuGlyAsnAapGlyCys 40
DB 159 CCAgGATCTCCAGCGCTGCGCTCTCAACTGCTTGGTAGGCTCTCGGCAACGATGGCTGC 218
OY 41 ThrArgLeuThrAapPheLysCysHisCysSerLysProGluLeuProGlyGlnIleThr 60
DB 219 ACTGCTTGACCGACTTCAGTGCACAGCTGCTCCAGGCTGAGCTCCAGAGACAGATCACT 278
OY 61 ProCysValGluGluAlaCysProLeuAsnAapAlaArgIleSerValSerAsnIleValVal 80
DB 279 CCTGCGTTGAGAGAGCGCTGCGCTCTCGACGCGCGATCTCCGTCCTCCAACTCGTCGT 338
OY 81 AapGlnCysSerLysAlaGlyValProIleAapIleProProValAspThrThrAlaAla 100
DB 339 GACCAgTCTCCAGGCGCGGTGTCCTCAATGGACATCCCAAGATTGACACCAAGCCCGCTT 398
OY 101 ProGluProSerGluThrAlaGluProThrAlaGluProThrGluGluProThrAlaGlu 120
DB 399 CCGAGGCAATCCGAGACCGCTGAGCCGCTGAGCCCAAGGAGAGAGCCCACTCGCGAG 458
OY 121 ProThrAlaGluProThrAlaGluProThrHisGluProThrGluGluProThrAlaVal 140
DB 459 CCTACCGCTGAGCCCAAGCGCTGAGCTCATGAGCCCAAGGAGAGAGCCCACTCGCGAG 518
OY 141 ProThrGlyThrGlyGlyValProThrGlyThrGlySerPheThrValThrGlyArg 160
DB 519 CCAACGGGACACTGGGGGTGTGTCCTGACCTGGACCGGTTCTTCACTGCTGAGCA 578
OY 161 ProThrAlaSerThrProAlaGluPheProGlyAlaGlySerAsnValArgAla-SerVal 180
DB 579 CCAACTGGCTCCACCCAGCTGAGTTCAGAGTGTGCTGCTCCAACTGCTGCTCCAGCT 638
OY 180 IGlyGlyIleAlaAlaAlaLeuLeuGlyLeuAlaAlaIleVal 194
DB 639 TGGCGGCAATGCTGCTGCTCTCGGTCTCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681
```

RESULT 2
BF252723

LOCUS BF252723 557 bp mRNA linear EST 15-NOV-2001
DEFINITION EST419986 Coccidioides immitis spherule cDNA library Coccidioides immitis cDNA clone CIABA21 5' sequence, mRNA sequence.

ACCESSION BF252723
VERSION BF252723.1 GI:16932866
KEYWORDS EST.
SOURCE Coccidioides immitis
ORGANISM Coccidioides immitis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Oryzales; mitosporic Oryzales; Coccidioides.

REFERENCE 1 (bases 1 to 557)
AUTHORS Gardner, M.J. and Kirkland, T.
TITLE Generation of ESTs from Coccidioides immitis spherule cDNA library
JOURNAL Unpublished
COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org

FEATURES
source Location/Qualifiers
1..557

/organism="Coccidioides immitis"
/mol_type="mRNA"
/db_xref="taxon:5501"
/clone="CIABA21"
/dev_stage="spherule"
/lab_host="SOLR"
/clone_lib="Coccidioides immitis spherule cDNA library"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 110 a 197 c 117 g 133 t
ORIGIN

Alignment Scores:
Pred. No.: 2,21e-45 Length: 557
Score: 706.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 69.01% Indels: 0
DB: 10 Gaps: 0

US-10-081-935-4 (1-194) x BF252723 (1-557)

```
OY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu 20
DB 161 ATGCAgTTCCTCAGCGCTCTCATGCTCGTGGCGGCGCTCGCAGATCCAGGCTC 220
OY 21 ProAapIleProProCysAlaLeuAsnCysPheValGluAlaLeuGlyAsnAapGlyCys 40
DB 221 CCAgGATCTCCAGGCGCGGTGTCCTCAATGGACATCCCAAGATTGACACCAAGCCCGCTT 280
OY 41 ThrArgLeuThrAapPheLysCysHisCysSerLysProGluLeuProGlyGlnIleThr 60
DB 281 ACTGCTTGACCGACTTCAGTGCACAGCTGCTCCAGGCTGAGCTCCAGAGACAGATCACT 340
OY 61 ProCysValGluGluAlaCysProLeuAsnAapAlaArgIleSerValSerAsnIleValVal 80
DB 341 CCTGCGTTGAGAGAGCGCTGCGCTCTCGACGCGCGATCTCCGTCCTCCAACTCGTCGT 400
OY 81 AapGlnCysSerLysAlaGlyValProIleAapIleProProValAspThrThrAlaAla 100
DB 401 GACCAgTCTCCAGGCGCGGTGTCCTCAATGGACATCCCAAGATTGACACCAAGCCCGCTT 460
OY 101 ProGluProSerGluThrAlaGluProThrAlaGluProThrHisGluProThrGluGlu 120
DB 461 CCGAGGCAATCCGAGACCGCTGAGCCCAAGCTGAGCCCAAGGAGAGAGCCCACTGCGAG 520
OY 121 ProThrAlaGluProThrAlaGluProThrHisGlu 132
```

DB		521	CTTACCGCTGACCCACCGCTGACCGACTATGAG	556
RESULT 3				
LOCUS	BF251334			
DEFINITION	EST418594 Coccidioides immitis spherule cDNA library Coccidioides immitis cDNA clone CIAF68 5' sequence, mRNA sequence.	531 bp	mRNA	linear
ACCESSION	BF251334			EST 15-NOV-2001
VERSION	BF251334.1	GI:16931477		
KEYWORDS	EST.			
SOURCE	Coccidioides immitis			
ORGANISM	Coccidioides immitis			
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onygenales; mitosporic Onygenales; Coccidioides.			
AUTHORS	1 (bases 1 to 531)			
TITLE	Gardner,M.J. and Kirkland,T.			
JOURNAL	Generation of ESTs from Coccidioides immitis spherule cDNA library Unpublished			
COMMENT	Contact: Malcolm J. Gardner Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301 838 3519 Fax: 301 838 0208 Email: gardner@tigr.org. Location/Qualifiers			
FEATURES	1..531			
source	/organism="Coccidioides immitis"			
	/mol_type="mRNA"			
	/db_xref="taxon:5501"			
	/clone="CIAF68"			
	/dev_stage="spherule"			
	/lab_host="SOLR"			
	/clone_job="Coccidioides immitis spherule cDNA library"			
	/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI"			
BASE COUNT	107 a	190 c	115 g	119 t
ORIGIN				
Alignment Scores:				
Pred. No.:	5..11e-45	Length:	531	
Score:	701.00	Matches:	131	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	68.52%	Indels:	0	
DB:	10	Gaps:	0	
US-10-081-935-4 (1-194) x BF251334 (1-531)				
OY	1	MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu	20	
DB	137	ATGCAGTTCCTCAGCGCTCTCATCGCTCTCGCGCTGCCGCTCCGACGATGCCAGCTC	196	
OY	21	ProAspIleProProCysAlaLeuAsnGlyPheValGluAlaLeuGlyAsnAspGlyCys	40	
DB	197	CCAGACATCCACCTTGCGCTCTCACTCGCTTCCTGTAAGCTCTTCGCAACATGCGCTGC	256	
OY	41	ThrArgLeuThrAspPheLeuGlyValProIleAspIleProGluLeuProGlyGlnIleThr	60	
DB	257	ACTCGCTTGACCGACTTCAGTGCACATCGCTCCAAAGCTGAGCTCCGACGATGACTACT	316	
OY	61	ProCysValGluGluAlaCysProLeuAspAlaArgIleSerValSerAsnIleValVal	80	
DB	317	CTTGCGCTGAGAGGCTGCGCTCTCGACGCCGCTATCTCGCTCCACATGATGCTGTT	376	
OY	81	AspGlnCysSerLeuAlaGlyValProIleAspIleProProValAspThrThrAlaAla	100	
DB	377	GACCACTGCTCCAAAGCGCGTGTCCCAATTGACATCCACCAAGTTCACACCGCGCT	436	
OY	101	ProGluProSerGluThrAlaGluProThrAlaGluProThrGluGluProThrAlaGlu	120	
DB	437	CCCGAGCCATCGAGACCGCTGAGCCACCGCTGAGCCAAACGAGAGAGCCACGTCGAG	496	

QY	121	ProthralagluProthralagluProthrhis	131
Bf251817			
Db	497	CTACCGCTGAGCCACCGCTGAGCGACTCAT	529
RESULT 4			
LOCUS	Bf251817	395 bp	mRNA
DEFINITION	EST419079 Coccidioides immitis spherule cDNA library Coccidioides immitis cDNA clone CIAAM80 5' sequence, mRNA sequence.	linear	EST 15-NOV-2000
ACCESSION	Bf251817		
VERSION	Bf251817.1	GI:16931960	
KEYWORDS	EST.		
SOURCE	Coccidioides immitis		
ORGANISM	Coccidioides immitis		
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onygenales; mitosporic Onygenales; Coccidioides.		
AUTHORS	1 (bases 1 to 395)		
TITLE	Gardner,M.J. and Kirkland,T.		
JOURNAL	Generation of ESTs from Coccidioides immitis spherule cDNA library		
COMMENT	Unpublished		
CONTACT	Malcolm J. Gardner		
DEPARTMENT	Department of Eukaryotic Genomics		
INSTITUTE	The Institute for Genomic Research		
ADDRESS	9712 Medical Center Drive, Rockville, MD 20850, USA		
TELEPHONE	Tel: 301 838 3519		
FAX	Fax: 301 838 0208		
EMAIL	Email: gardner@tigr.org		
FEATURES	Location/Qualifiers		
SOURCE	1..395		
	/organism="Coccidioides immitis"		
	/mol_type="mRNA"		
	/db_xref="taxon:5501"		
	/clone="CIAAM80"		
	/dev_stage="spherule"		
	/lab_host="SOLR"		
	/clone_1lb="Coccidioides immitis spherule cDNA library"		
	/note="Vector: pbluescript SK(-); Site_1: EcoR1; Site_2: Xho1"		
BASE COUNT	68 a 158 c 96 g 73 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	6,62e-41	Length:	395
Score:	646.00	Matches:	125
Percent Similarity:	98.43%	Conservative:	0
Best Local Similarity:	98.43%	Mismatches:	2
Query Match:	63.15%	Indels:	0
DB:	10	Gaps:	0
US-10-081-935-4 (1-194) x Bf251817 (1-395)			
QY	68	ProleuaapaalarglleserValserAenllevalaAapginCyserryalaGly	87
Db	2	CCTCGAGCGCCGATATCGTCCCAACATCGCTTGACAGTCTCCAAAGCCGCT	61
QY	88	ValProilleapilleProProvalAasphrthrAlaAaPProGluProSerGluThrAla	107
Db	62	GTCCCAATGACATCCACAGATTGACACCCGCCGCTCCGAGCCATCGAGACCGCT	121
QY	108	GlupProthralagluProthGluGluProthralagluProthralagluProthrala	127
Db	122	GAGCCACCGCTGAGCCACCGAGAGAGCCACTGCCAGCTTACCGCTGAGACCCAGCT	181
QY	128	GlupProthrhiegluProthGluGluProthralaValProthGlyThGlyGly	147
Db	182	GAGCGAGATCATGAGCCACCGAGAGAGCCACTGCTGCCAACCGGCACTGGCGTGGT	241
QY	148	ValProthrhGlyThGlyserPhethThValThGlyArgPProthralaserThProAla	167
Db	242	GTCCCACTGGACACCGGTTCCTACCTCATCGACATCGACACCACTGCTCCACCCAGCT	301
QY	168	GlupheProglalaglyserAasvalArgAlaserValGlyGlylleAlaAlaAla	187

|||||
Db 302 GAGTTCACAGTGGCTGCTCCACGCTCCGTGCGACAGCTTGCGGCGCATTTGCTGCTCTC 361
OY 188 LeuGlyLeuAlaAlaTyrLeu 194
|||||
Db 362 CTCGGGCTCGCTGCTACCTG 382

RESULT 5
BF252856 401 bp mRNA linear EST 15-NOV-2001
LOCUS BF252856
DEFINITION Coccidioides immitis spherule cDNA library Coccidioides
immitis cDNA clone CIAB70 5' sequence, mRNA sequence.
ACCESSION BF252856
VERSION BF252856.1 GI:16932999
KEYWORDS EST.
SOURCE Coccidioides immitis
ORGANISM Coccidioides immitis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE 1 (bases 1 to 401)
AUTHORS Gardner, M.J. and Kirkland, T.
TITLE Generation of ESTs from Coccidioides immitis spherule cDNA library
JOURNAL Unpublished
COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@cigr.org.
Location/Qualifiers
source 1..401
/organism="Coccidioides immitis"
/mol_type="mRNA"
/db_xref="taxon:5501"
/clone="CIAB70"
/dev_stage="spherule"
/lab_host="SOLR"
/clone_lib="Coccidioides immitis spherule cDNA library"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 77 a 149 c 89 g 86 t
ORIGIN
Alignment Scores:
Pred. No.: 5.86e-38 Length: 401
Score: 608.00 Matches: 114
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 59.43% Indels: 0
DB: 10 Gaps: 0

US-10-081-935-4 (1-194) x BF252856 (1-401)

OY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu 20
Db 59 ATGAGTTCTCTACGCTCTCATCGCTCTGCTGCGGCTGCGAGTGGCCAGCTTC 118
OY 21 ProAspIleProProCysAlaLeuAsnCysPheValGluAlaLeuGlyAsnAspGlyCys 40
Db 119 CCAGACATCCACACCTTGCGCTCTCAACTGCTTGAGGCTCTCGGCAACGATGGCTGC 178
OY 41 ThrArgLeuThrArgPheLeuPheLeuGlyCysSerLeuProGluLeuProGlyGlnIleThr 60
Db 179 ACTGCTTGACCGACTTCAGTGCACCTGCTCCAAAGCTGAGCTCCCGAGACAGTCCCT 238
OY 61 ProCysValGluGluAlaCysProLeuAspAlaArgIleSerValSerAsnIleValVal 80
Db 239 CTTGGCGTTGAGAGGCGCTGCTCTCGACGCCGATCTCGTCTCAACATGCTGCTT 298
OY 81 AspGlnCysSerLeuAlaGlyValProIleAspIleProProValAspThrThrAlaAla 100
Db 299 GACCAGTGCTCAAGGCGGCTGCTCCCAATTGACATCCACACGATTGACACACGCGCT 358

OY 101 ProGluProSerGluThrAlaGluProThrAlaGluProThr 114
Db 359 CCCAGGCCATCCGAGACCGCTGAGCCACCGCTGAGCCAC 400

RESULT 6
BF251279 395 bp mRNA linear EST 15-NOV-2001
LOCUS BF251279
DEFINITION Coccidioides immitis spherule cDNA library Coccidioides
immitis cDNA clone CIAAFl1 5' sequence, mRNA sequence.
ACCESSION BF251279
VERSION BF251279.1 GI:16931422
KEYWORDS EST.
SOURCE Coccidioides immitis
ORGANISM Coccidioides immitis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE 1 (bases 1 to 395)
AUTHORS Gardner, M.J. and Kirkland, T.
TITLE Generation of ESTs from Coccidioides immitis spherule cDNA library
JOURNAL Unpublished
COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@cigr.org.
Location/Qualifiers
source 1..395
/organism="Coccidioides immitis"
/mol_type="mRNA"
/db_xref="taxon:5501"
/clone="CIAAFl1"
/dev_stage="spherule"
/lab_host="SOLR"
/clone_lib="Coccidioides immitis spherule cDNA library"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 78 a 133 c 78 g 106 t
ORIGIN
Alignment Scores:
Pred. No.: 8.03e-26 Length: 395
Score: 451.00 Matches: 84
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 44.09% Indels: 0
DB: 10 Gaps: 0

US-10-081-935-4 (1-194) x BF251279 (1-395)

OY 1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu 20
Db 142 ATGAGTTCTCTACGCTCTCATCGCTCTGCTGCGGCTGCGAGTGGCCAGCTTC 201
OY 21 ProAspIleProProCysAlaLeuAsnCysPheValGluAlaLeuGlyAsnAspGlyCys 40
Db 202 CCAGACATCCACACCTTGCGCTCTCAACTGCTTGAGGCTCTCGGCAACGATGGCTGC 261
OY 41 ThrArgLeuThrArgPheLeuPheLeuGlyCysSerLeuProGluLeuProGlyGlnIleThr 60
Db 262 ACTGCTTGACCGACTTCAGTGCACCTGCTCCAAAGCTGAGCTCCCGAGACAGTCCCT 321
OY 61 ProCysValGluGluAlaCysProLeuAspAlaArgIleSerValSerAsnIleValVal 80
Db 322 CTTGGCGTTGAGAGGCGCTGCTCTCGACGCCGATCTCGTCTCAACATGCTGCTT 381
OY 81 AspGlnCysSer 84
Db 382 GACCAGTGCTCC 393

RESULT 7

LOCUS	BQ767421	602 bp	mRNA	linear	EST 26-JUL-2002
DEFINITION	EBR008_SQ008_122_R root, 3 week, drought-stressed, cv Optic, EBR008 Hordeum vulgare subsp. vulgare cDNA clone EBR008_SQ008_122 5', mRNA sequence.				
ACCESSION	BQ767421				
VERSION	BQ767421.1	GI:21975895			
KEYWORDS	EST.				
SOURCE	Hordeum vulgare subsp. vulgare				
ORGANISM	Hordeum vulgare subsp. vulgare				
REFERENCE	Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophytes, Magnoliophyta, Liliopsida, Poales, Poaceae, Pooidaeae, Triticeae, Hordeum.				
AUTHORS	1 (bases 1 to 602) Hedley, P., Iku-Hira, G., Caldwell, D., McCallum, N., Mudie, S., Cardle, L., Ramsey, L., Machray, G., Marshall, D.F.M. and Waugh, R.				
TITLE	Development of Barley Transcriptome Resources				
JOURNAL	Unpublished				
COMMENT	Contact: Waugh R, Marshall DF Genome Dynamics/Computational Biology Scottish Crop Research Institute Invergowrie, Dundee, DD2 5DA, Scotland, UK Tel: 00 44 1382 562731 Fax: 00 44 1382 562426 Email: est@scri.sari.ac.uk. Location/Qualifiers				
FEATURES	source				
	1..602				
	/organism="Hordeum vulgare subsp. vulgare"				
	/mol_type="mRNA"				
	/culivar="Optic"				
	/db_xref="taxon:112509"				
	/clone="EBR008_SQ008_122"				
	/issue_type="root"				
	/dev_stage="3 week"				
	/lab_host="DHL08"				
	/clone_lib="root, 3 week, drought-stressed, cv Optic, EBR008"				
	/note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I; Non-normalised library, directionally cloned into pSPORT1. Derived from roots of 3 week old drought stressed barley plants. Developed as part of the barley transcriptome resources of BBSRC/SEERAD funded cereal IGF (Investigating Gene Function) project."				
BASE COUNT	100 a 251 c 129 g 122 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	1.22e-22	Length:	602		
Score:	412.50	Matches:	84		
Percent Similarity:	59.8%	Conservative:	25		
Best Local Similarity:	46.1%	Mismatches:	56		
Query Match:	40.3%	Indels:	17		
DB:	13	Gaps:	1		
US-10-081-935-4 (1-194) x BQ767421 (1-602)					
OY	1	Merc1npheserH1sAla1eu1lAla1euVal1aAgl1y1Leu1a1aSer1a1Gln1eu	20		
DB	45	ATGAAGTTCACACCACTGATGATGCTTCGTGTCGGCGGCTTCGCGACGCCACCTG	104		
OY	21	ProAsp1leProProCySA1aLeuAsnCySPheVal1G1ua1aLeuG1yAsnAspG1yCyS	40		
DB	105	CCCGATGCGCCGCGCTGCTGCTAGCGTCTTCCTCTGCTGCTTCGACAGTGAATGCTTGC	164		
OY	41	Th1rAgl1euTh1rAsp1le1ySCyH1sCySPSer1ySPG1leuP1roG1yGln1leTh1r	60		
DB	165	TCCACAGTGTCCGACCTTCACCTGCACCTGCACGACCCGAGCGCTCGTAGACAGATCACT	224		
OY	61	ProCySA1aG1uG1ua1aCySPProLeuAspAla1rG1leSerVal1SerAsn1leVal1	80		
OY	225	CTTGTGTTCGAAGACCTGCGAGCTCAAGACCAATTCCTGCTTCACGCCGCTGCTC	284		
OY	81	AspG1nCySer1ySA1aG1yVal1Pro1leAsp1leProProVal1AspTh1rTh1a1a	100		

Db		265	GGCCAGTCTCTTCGCGCCGCACACCCTACTTCGATTTCCCCTCATGACACCAAGCCCTTCC	344
Oy		101	ProglutPogserGIuthrAlaGlupProthrAlaGlupProthrAlaGlupProthrAla	120
Db		345	TCCAGCGCCTCCAGGGCTCCAGACACTCGAGCTCCGCCCGCTGAGACAGCCGCCAG	404
Oy		121	ProthrAlaGlupProthrAlaGlupProthrAlaGlupProthrAlaGlupProthrAla	140
Db		405	ACCACCGCTCCACACGAGAGACTCCAGCCCGTTAGTCTCTCTCTCTCCCTGT	464
Oy		141	ProthrGIY-----	143
Db		465	GAGACCGGACCTCCACACCGGTGCGCAGTAATACTACTACCGTGTTTCCACCAAGTC	524
Oy		144	ThrglyGIyGIyValProthrGIYThrglyseRhetThrValThnglyValProthrAla	163
Db		525	AGCTCTCCGGTAGTCTCCACCTCCAGCCCGCTGAGACCACTACCGCTCACCACTGCC	584
Oy		164	SerThr	165
Db		585	TCTACC	590
RESULT 8				
LOCUS	BQ767128	538 bp	mRNA	linear EST 26-JUL-2002
DEFINITION	EBro08_S0008_E05_R_roof_3 week, drought-stressed, cv Optic, EBro08_S0008_E05_5', mRNA			
ACCESSION	BQ767128			
VERSION	BQ767128.1	GI:21975602		
KEYWORDS	Hordeum vulgare subsp. vulgare cDNA clone			
SOURCE	EST.			
ORGANISM	Hordeum vulgare subsp. vulgare			
	Hordeum vulgare subsp. vulgare			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae			
	; Triticeae; Hordeum.			
	1 (bases 1 to 538)			
REFERENCE	Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardie,L.,			
AUTHORS	Ramsay,L., Machray,G., Marshall,D.F.M. and Maugh,R.			
TITLE	Development of Barley Transcriptome Resources			
JOURNAL	Unpublished			
COMMENT	Contact: Maugh R, Marshall DP			
	Genome Dynamics/Computational Biology			
	Scottish Crop Research Institute			
	Invergowrie, Dundee, DD2 5DA, Scotland, UK			
	Tel: 00 44 1382 562731			
	Fax: 00 44 1382 562426			
	Email: est@scri.bari.ac.uk.			
FEATURES	Location/Qualifiers			
source	1..538			
	/organism="Hordeum vulgare subsp. vulgare"			
	/mol_type="mRNA"			
	/culivar="Optic"			
	/db_xref="taxon:112509"			
	/clone="EBro08_S0008_E05"			
	/tissue_type="root"			
	/dev_stage="3 week"			
	/lab_host="DH10B"			
	/clone_id="root_3 week, drought-stressed, cv Optic,			
	EBro08"			
	/note="Vector: pSPORT1, Site_1: Sal I; Site_2: Not I;			
	Non-normalised library, directionally cloned into pSPORT1.			
	Derived from roots of 3 week old drought stressed barley			
	plants. Developed as part of the barley transcriptome			
	resources of BBSRC/SERAD funded cereal IGF (Investigating			
	Gene Function) project."			
BASE COUNT	91 a 217 c 119 g 111 t			
ORIGIN				

TITLE		The Paracoccidioides brasiliensis EST genome project	
JOURNAL		Eukaryotic Cell, (2002) In press	
COMMENT		Contact: Gustavo Henrique Goldman Laboratory of Molecular Biology Universidade de Sao Paulo - USP - FCRP Av do Cafe S/N, CEP: 14040-903, Ribeirao Preto - SP, Brazil Email: ggoldman@usp.br	
FEATURES		Location/Qualifiers	
source		1..311	
BASE COUNT		73 A 88 C 59 G 91 T	
ORIGIN			
Alignment Scores:			
Pred. No.:		5,87e-16	
Score:		322.00	
Percent Similarity:		93.06%	
Best Local Similarity:		77.78%	
Query Match:		31.48%	
DB:		13	
US-10-081-935-4 (1-194) x BQ499065 (1-311)			
OY	1	MeGlnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu	20
DB	91	ATGCAAGTCTCTCAGCGCTCTCATCTCTGTTGGCGCCAGCTACCAATCTCAGCTC	150
OY	21	ProhApIleProPtoCysAlaLeuAsnCysPheValGluAlaLeuGlyAsnAspGlyCys	40
DB	151	CCTGATATCCACCACTTGGCGCTTGAAGCTTTCCTGTAATCTCTCGAAACACCGCGTGC	210
OY	41	ThrArgLeuThrAspPheLeuSerCysHisCysSerLeuPProGluLeuProGlyGlnIleThr	60
DB	211	TCCCAAGCTGACCGAATTCGAGTCCACATCGCAAAAGCCTGAGTTCCTCGAAGATCACT	270
OY	61	ProCysValGluGluAlaCysPProLeuAspAlaArg	72
DB	271	CCTGCGTGGAGAAATCTTGGACGCTGCTGCCCA	306
RESULT 11			
LOCUS	BM077163	667 bp	mRNA linear EST 05-FEB-2002
DEFINITION	TrEST-A0723 TrEST-A Hypocrea jecorina cDNA clone Tr-A0723 5', mRNA		
ACCESSION	BM077163		
VERSION	BM077163.1		
KEYWORDS	EST.		
SOURCE	GI:18498345		
ORGANISM	Hypocrea jecorina		
REFERENCE	Hypocrea jecorina		
AUTHORS	Eukaryotic Cell, (2002) In press		
TITLE	1 (bases 1 to 667)		
JOURNAL	Chamberg, F.S., Bonaccorsi, E.D., Ferreira, A.J.S., Ramos, A.S.P.,		
MEDLINE	Ferreira, J.R., Jr., Abrahao-Neto, J., Farah, J.P.S., and El-Dor, H.		
PUBMED	Elucidation of the metabolic fate of glucose in the filamentous		
COMMENT	Fungal Trichoderma reesei using expressed sequence tag (EST)		
CONTACT	analysis and cDNA microarrays		
CONTACT	J. Biol. Chem. 277 (16), 13983-13988 (2002)		
CONTACT	11825887		
CONTACT	11825887		
CONTACT	Contact: El-Dor, Hama		
CONTACT	Department of Biochemistry		
CONTACT	Institute of Chemistry, University of Sao Paulo		
CONTACT	Avenida Professor Lineu Prestes, 748, Sao Paulo, SP, 05508-900,		
CONTACT	BRASIL		
CONTACT	Tel: (55) 11-38183848		
CONTACT	Fax: (55) 11-38183848		
CONTACT	Email: dorry@iq.usp.br		
CONTACT	PCR Primers		
CONTACT	FORWARD: Universal M13 forward primer		

<p>BACKWARD: Universal M13 reverse primer Plate: 8 row: E column: 3 Seq primer: M13 reverse primer High quality sequence stop: 667 POLYA=No.</p>	
FEATURES	<p>Location/Qualifiers</p> <p>1..667</p> <p> /organism="Hyphocrea jecorina"</p> <p> /mol_type="mRNA"</p> <p> /strain="QM9414 (ATCC26921)"</p> <p> /db_xref="taxon:51453"</p> <p> /clone_lib="Tr-A0723"</p> <p> /sex="Asexual"</p> <p> /tissue_type="Mycelia"</p> <p> /dev_stage="18 hr glycerol-grown culture"</p> <p> /lab_host="E. coli S0LR cells (kanamycin resistant)"</p> <p> /clone_lib="TREST-A"</p> <p> /notes="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2: XhoI; anamorph=Trichoderma reesei; Cloned unidirectional, 5' end of the cDNA cloned into EcoRI site of pBluescript. Primer: Oligo (dT). Average insert size: 1, kb. Uni-ZAP XR Vector system -5' adaptor sequence: 5'GAATTCGGCAGAG3' -3' adaptor sequence: 5'CTCAGTCTTTTCTTTT3"</p>
BASE COUNT	<p>98 a 266 c 188 g 115 t</p>
ORIGIN	
Alignment Scores:	
Pred. No.:	1.16e-07 Length: 667
Score:	219.50 Matches: 72
Percent Similarity:	44.19% Conservative: 23
Best Local Similarity:	33.49% Mismatches: 69
Query Match:	21.46% Indels: 51
DB:	12 Gaps: 14
US-10-081-935-4 (1-194) x BM077163 (1-667)	
QY	23 11eProProCySa1a1euaBncysPheVa1G1ua1a1eug1A5naBp01CYeThrArg 42
DB	33 GTCCCTCATGGCCCAAGCGTCTCTCTGCTGGCCGCCCGGGCGTGGCTGCAC--- 89
QY	43 LeuThrAspPhe1scYsh1scYSer1sPProGluLeuProGlu1Gln1leThr----- 60
DB	90 GTCACGACTTTCACATGCGACATGC---AAGCCCGAGGCTCAC;ACAGCCTGACCATCTG 141
QY	61 -----ProCySa1a1GluGlu1a1CySPro-----LeuAsp1a1Arg1leSerVa1 75
DB	147 CTGCTGCTCGCTCGCTCCCACTGCTGCTGCTGCTGCTCTCTCTCAGAGCGGTATCACCGCA 201
QY	76 SerAsn11eVa1a1aBpG1ncYser1sA1aG1Va1Pro1eAsp11eProProVa1 95
DB	207 GCCAGCTCAAGTTCG---GCTGTGCGACGCGCTGCGCCACGCTGCTGCGCTGGAAT 261
QY	96 AspThrThThAla1a1aProGluPro-----SerGluThThAlaGlu---ProThr 111
DB	264 TGCACAGCAGCAGCGCTCTCTCCCGCGCGAGCTCACAGCACCCGGAATGCCCGCAGC 322
QY	111 Alag1uProThrg1uGluProThThAlaGluPro-----ThThAla 122
DB	324 AGCTCCCGGTGACGAGCCCTCCGGAAGCCGCTCTCTGCTTCTTCTTCTTCTCT 381
QY	124 GluProThThAlaGluProThTh1sGluProThr----- 133
DB	384 CCCCAGCAGCAGCGCTACAGCAGACCCATCGCGCGCAACGAGAAAGCAGCCGCT 441
QY	135 -----GluGluProThThAla-----Va1ProThrg1YThG1y 144
DB	444 GAGCAAGCGCGGACACAGCCCAACAGCGGATTCACACATGCCCCAGCGGCTT---GGC 500
QY	147 GlyVa1ProThrg1YThrg1SerPheThThVa1Thrg1YArgProThThAlaSerThPro 166
DB	501 TTCGACCCGACCGGACCTGGTTCCGAGACGACTACTGAGAGAGAGAGCGAAGCCAGCC 561

QY 167 AlaGlu-----PheProGlyAlaGlySerValAlaGlyAlaSerVal----- 180
 Db 561 ACTGGCGCCGACGACTCTTCCCCGAGCGCGGCTTCCAGTCGAGATGACCTCGTGTCT 620
 QY 181 GlyGlyIleAla-----AlaIleuLeuGlyLeu 190
 Db 621 TGTGCTTTTGGCGCTACTGCGGCGCATTCAGCTTTGCTGGATTC 665

RESULT 12
 BM866652 564 bp mRNA linear EST 06-MAY-2003
 LOCUS mgc8006x809f.b Magnaporthe grisea CS uni-zap XR library Magnaporthe
 DEFINITION grisea cDNA clone mgc8006x809 5', mRNA sequence.
 ACCESSION BM866652.1 GI:19234334
 VERSION 1
 KEYWORDS
 SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
 ORGANISM
 REFERENCE
 AUTHORS Ebbole,D.J., Yuan,J., Thomas,T.L., Bobrowicz,P., Lu,G., Bhatteai
 TITLE Expressed sequence tags from the rice blast fungus, Magnaporthe
 grisea
 JOURNAL
 COMMENT Unpublished
 Contact: Ebbole DJ
 Department of Plant Pathology & Microbiology
 Texas A&M University
 Peterson Bldg, MS2132, College Station, TX 77843-2132, USA
 Tel: 979 845 4831
 Fax: 979 845 6483
 Email: d-ebbole@tamu.edu
 Chromatogram file of this sequence is available, see contact person
 ;Best nr hit (April. 22, 2003) gb|EAA29323.1| predicted protein
 [Neurospora crassa] 55 6e-07
 PCR primers
 FORWARD: T3 primer
 BACKWARD: T7 primer
 Plate: mgc8006 row: B column: 09
 Seq primer: T3.

FEATURES
 source
 1..564
 Location/Qualifiers
 /organism="Magnaporthe grisea"
 /mol_type="mRNA"
 /strain="Guy11"
 /db_xref="taxon:148305"
 /clone="mgc8006x809"
 /sex="Mat1-2 hermaphrodite"
 /cell_type="conidia"
 /clone_1fb="Magnaporthe grisea CS Uni-Zap XR Library"
 /note="Vector: pBluescriptSK-; Site_1: EcoRI; Site_2: XhoI
 ; Unidirectional cloning. EcoRI side has T3 primer and
 predominantly 5' reads. T7 primer on XhoI side of insert.
 Conidial library. Point inoculation of Guy11 at center of
 oatmeal agar plate. Conidia were harvested after two weeks
 of growth. Sequences were processed by one of two methods.
 Where a full-length alignment to the M. grisea genome
 sequence was available, the EST sequence was trimmed
 according to the alignment, otherwise sequence quality was
 assessed using phredPhrap version 991019 and trimmed
 according to phd files (0.05) and for vector segs."

BASE COUNT 100 a 242 c 128 g 94 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,98e-06 Length: 564
 Score: 202.50 Matches: 64
 Percent Similarity: 47.42% Conservative: 28
 Best Local Similarity: 32.95% Mismatches: 81
 Query Match: 19.79% Indels: 22
 DB: 12 Gaps: 10

US-10-081-935-4 (1-194) x BM866652 (1-564)

QY 1 MetGlnPheSerHnIaIaIeuiIaIaIeuiValaIaGlyIleuAlaSerValaGlnIleu 20
 Db 33 ATGAAGTTCTCC---ATCATCTCTCGCGCGCGCGCCGACGCGCTGCTGCTGACGACTC 89
 QY 21 ProAspIleProProGlyValaIleuAnCyPheValaGluAlaIleuGlyAsnAspGlyCys 40
 Db 90 TCTTCCTCCCGCGCTGCGCTGTCGACGTGTGAGCAGCGCTATTGCTTCCGCTGC 149
 QY 41 ThrArgLeuThrAspPheIaCyPheIaSerIa-----SerIaProGluLeuProGlyIle 58
 Db 150 AACGCG---ACCGACGTGCTCGCTCCGACGTGTGTGAGAGAGAGACATTGCCAACGTC 206
 QY 59 IleThrProGlyValaGluIaIaCyPheIleuAspAlaArgIleSerValSerAnIle 78
 Db 207 GCTACTCTCTGCTGCTGCTGCGACGACGACCGCTGCGACCTCTCCAGGCTGCCACC 266
 QY 79 -ValValaAspGlnIaSerIaValaGlyValaProIleAspIleProValaAspThrTh 98
 Db 267 AACGCGAGCGCTGTGCMAGACGTCCCGACCGCTGTTCTTACTCTCT-----CC 317
 QY 98 AlaAlaIaProGluProSerGluThrAlaGluProThrAlaGluProThrGluGluProTh 118
 Db 318 AGCGCGCCCTTACCAACACCCACC---CCGCAACTTTCGCGCTGCCGCGCACGACCC 374
 QY 118 AlaGluProThrAlaGluProThrAlaGluProThrIleGluProThrGluGluProTh 138
 Db 375 ATGCCCTCATGTGATGATGCGTGTGCCCAAGACCAACGCG-ACCTCTGCTGCTCTCCAC 433
 QY 138 AlaValProThrGluThrGlyGlyGlyVala-----ProThrGlyThrGlySerPheTh 156
 Db 434 CGCTTCCCGCCACGACGACGCGGCGAGAGAACCCAGCGGACACGCGC----- 488
 QY 156 rValThrGlyIaProThrAlaSerThrProAlaGluIaPheProGlyAlaGlySerAsnVa 176
 Db 489 -----GCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 532

QY 176 IArgIaSerValaGlyIleAlaIaIaIeuiGly 189
 Db 533 C---GCCACTATCGCG-----GCTGCTGCTCTTCTTGT 563

RESULT 13
 BM866040 553 bp mRNA linear EST 06-MAY-2003
 LOCUS mgc8004x016f.b Magnaporthe grisea CS uni-zap XR library Magnaporthe
 DEFINITION grisea cDNA clone mgc8004x016 5', mRNA sequence.
 ACCESSION BM866040
 VERSION 1
 KEYWORDS
 SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
 ORGANISM
 REFERENCE
 AUTHORS Ebbole,D.J., Yuan,J., Thomas,T.L., Bobrowicz,P., Lu,G., Bhatteai
 TITLE Expressed sequence tags from the rice blast fungus, Magnaporthe
 grisea
 JOURNAL
 COMMENT Unpublished
 Contact: Ebbole DJ
 Department of Plant Pathology & Microbiology
 Texas A&M University
 Peterson Bldg, MS2132, College Station, TX 77843-2132, USA
 Tel: 979 845 4831
 Fax: 979 845 6483
 Email: d-ebbole@tamu.edu
 Chromatogram file of this sequence is available, see contact person
 ;Best nr hit (April. 22, 2003) gb|EAA29323.1| predicted protein
 [Neurospora crassa] 51 8e-06
 PCR primers
 FORWARD: T3 primer
 BACKWARD: T7 primer

Db	656	TGCTTGACCGAAGCTTGCCAGGTGATGACTACTACCAAGGCCGAGAAAGCTGCGCGCAC	597
Oy	82	GlnCys-----SerLysAlaGlyValProIleasp	91
Db	596	GCCCTGCAAGGCTATGCTGCCACCGCTGCTCCGCTCCGACCTGACCGCTCCACCGACG	537
Oy	92	-----IleProValAspThrThrAlaIleProGluProSerGluThr	106
Db	536	AGTCTGGCTCTGTTACCGTCAGCATGACACCTTATCACTGGCTCCGCTAGCCTTCC	477
Oy	107	AlaGluProThrAlaGluProThrGluGluProThrAlaGluProThrAlaGlu-Proth	126
Db	476	GCCATCTTCTCAACATCCCTGATGAGAACCCATCCTCTGTTACTCCCGTCCGCGC	417
Oy	126	raIaGluProThrIstGluPro-----ThrGluProThrAlaIa	140
Db	416	AACGCAACCAATGACCAAGACATGACTAGAGAGTGCATCCGGAGGAGC-GCTGTAC	358
Oy	140	lProThrGlyThrGlyGlyValProThrGlyThrGlySerPheThrValThrGlyAr	160
Db	357	ACCCACGGGAACCTCTGGCGCG-----GGCGGCGAGCGGACACATCGCGCGTGC	307
Oy	160	gProThrAlaSerThrProAlaGluPheProGlyValaGlySerAsnValaGlaIaSerVa	180
Db	306	TAGCAGTGTCCCTACCGACGCT-----GGTCTGTACTCTGTGCGCCCTGTCGCAT	256
Oy	180	IGly-----GlyIleAlaAlaLeuLeuGlyIleu	190
Db	255	GGCTCTCTCCGCCATTATTCGTCGTCGATCCCTCTG	219
RESULT 15			
BU066219		731 bp	mRNA linear EST 26-AUG-2002
LOCUS			
DEFINITION	Fgr 9_K19_T3 Nitrogen-starved mycelia	Gibberella zeae	cdNA, mRNA
ACCESSION	BU066219		
VERSION	BU066219.1	GI:22506508	
KEYWORDS	EST.		
SOURCE	Gibberella zeae		
ORGANISM	Gibberella zeae		
	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;		
	Hypocreales; Hypocreales; Nectriaceae; Gibberella.		
REFERENCE	Trail,P., Xu,J., R., San Miguel,P., Halgren,R.G. and Kistler,H.C.		
AUTHORS	Analysis of expressed sequence tags from Gibberella zeae (anamorph		
TITLE	Fusarium graminearum)		
JOURNAL	Fungal Genet. Biol. 38 (2), 187-197 (2003)		
MEDLINE	22508120		
PUBMED	12620255		
COMMENT	Contact: Frances Trail Department of Plant Biology Michigan State University East Lansing, MI 48824, USA Tel: 517 432 2939 Fax: 517 353 1926 Email: trail@msu.edu Plate: 9 row: K column: 19.		
FEATURES			
source	Location/Qualifiers		
	1..731		
	/organism="Gibberella zeae"		
	/mol_type="mRNA"		
	/strain="NRRL 31084"		
	/db_xref="taxon:5518"		
	/clone_id="Nitrogen-starved mycelia"		
	/note="Vector: Uni-ZapIII; Site_1: EcoRI; Site_2: XhoI"		
BASE COUNT	151 a 255 c 172 g 153 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	7,12e-05	Length:	731
Score:	184.00	Matches:	66
Percent Similarity:	41.51%	Conservative:	20
Best Local Similarity:	32.08%	Mismatches:	91

Query Match: 17..99% Indels: 34
DB: 13 Gaps: 8

US-10-081-935-4 (1-194) x BU066219 (1-731)

Oy 6 Alaleu1lealaeValua1alaagilyLeualaSerA1agn---LeuproAsp1Iepro 24
||| :::: ||| ||||||||| ||||| |||
Db 120 GCTACCGCTTTCCTCCGCTGGCAGGCCGCTGCCTGTCTCACGAGCACTCATGTGTAACC 179

Oy 25 ProCyAlaleuanaenCyPhevalaIGluAlaleuclyvnaenaspQicyStmrArgLeuthr 44
||| ||||| |:
Db 180 GAGTGCTCTCCAATATTCTTAAGAATTCATCCCACAAGCTGGACTTC--GACTTGAC 236

Oy 45 AspheLueCyShIScySer-----LyasProLueupProGIgynIethPro 61
::: ||| ||| :::: ||| |||||||||
Db 237 AACACCGCATGCCATGCGAGCGCTGTTCCAAACAAGCTCGTCCCATCATCATCTCC 296

Oy 62 Cyva1eIgncIualacyProLeuaspAlaarg1IeserValSerbenIIevala1asp 81
||| ||||| :||| :|||
Db 297 TGCTTGACCGAGAGCTCCAGGTGATGACCTACTCTAACGCCCCGAAGAGCTGCCCGCAC 356

Oy 82 Glncy-----serLySa1agiValproIlaasp 91
||| ||| |||
Db 357 GCCTGCAAGCCCTTAGCTGCCACCGCTGGCTCCGCTCCGCACTGCCCTCCACCGAG 416

Oy 92 -----IIeproValaspThrThrAla1aprogiuProSergUthr 106
||| ||| :|||
Db 417 AGTCTGGCTCTGTACCGTCAGACATGACACCTCTATCACTGGCTCCGTAAGCTTCCC 476

Oy 107 AlaaguProthrAlaGu1ProthrgLuGu1ProthralaGu1Prothra1aGu1-Proth 126
||| :||| ||| ||| |||
Db 477 GCCATCTTCTCAACATCTCCGTATGATAAGCCCCATCCCTCTGTACTCCCGCTCCGC 536

Oy 126 rAlaGu1ProthrhiagiUp-----ThrgLuGu1Prothra1ava 140
||| ||| |||
Db 537 AAGCAACCATAGACCACCAAGACCATATCTAGAGAGCTGATCCGAGAGAGC-GCTGTAC 595

Oy 140 lProthrgLyThrsiglygyValProthrgLyThrgYserPherThervalThrglyar 160
||| ||| ||| ||| |||
Db 596 ACCCAGCGGAACCTCTGGCGCG-----GGCGCGCGCGCGCCACCATCGCGGTGC 646

Oy 160 gProthralaserhProhiagiUpheProglAlaGlYseranua1ArGalaserVa 180
||| ||| ||| ||| ||| ||| ||| |||
Db 647 TAGCAGTGCTCCATACGACGCT-----GGTgcgtGhAcTcctGcgccTcgtccat 697

Oy 180 IGlyGyllea1aa1aalaeuleugilyLeua1a 191
:||| :||| ||| ||| ||| |||
Db 698 GGCTCTCTCCCATATATGCTGCTGCTGATGCT 731

RESULT 16
BM872390 688 bp mRNA linear EST 06-MAY-2003
LOCUS LOCUS
DEFINITION mgmy002xP06.f.b Magnaporthe grisea my Uni-Zap XR library Magnaporthe
grisea cDNA clone mgmy002xP06 5', mRNA sequence.
VERSION BM872390
KEYWORDS BM872390.2 GI:30406827
SOURCE EST.
ORGANISM Magnaporthe grisea (anamorph: Pyricularia grisea)
MagnaPorte grisea
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
REFERENCE 1 (bases 1 to 688)
AUTHORS Ebbole,D.J., Yuan,J., Thomas,T.L., Bobrowsicz,P., Lu,G., Bhatterai,
'K. and Dean,R.A.
TITLE Expressed sequence tags from the rice blast fungus, Magnaporthe
grisea
JOURNAL Unpublished
COMMENT On Mar 7, 2002 this sequence version replaced gi:19240072.
Contact: Ebbole DJ
Department of Plant Pathology & Microbiology
Texas A&M University
Peterson Bldg, MS2332, College Station, TX 77843-2132, USA
Tel: 979 845 4831

Db 390 CCGTCCAGCTTCATCCAGCCCGGAGCTGACTGCTTCCACACACCGTCCGAGCC 449
Oy 113 ProthrgluGluProthralaGluProthralaGluProthralaGluProthriaglu 132
Db 450 CCGACGGAGACACCTCTGCGCGAGCCCTGCGGGGGCCGACCGCGCAACATGCGC 509
Oy 133 ProthrgluGluProthralaValProthrglyGlyGlyValPro----- 149
Db 510 GCTGCGAG-----TGACAGCATCTGGTGTGTGCTGCGGCTTGGGATGACC 563
Oy 150 ThrglyThrGlySerPheThrValThrGlyArg 160
Db 564 CTGGGCTTGAGCATGACAAACTCGGGCGCA 596

RESULT 19
BU383609/c BU383609 612 bp mRNA linear EST 08-MAR-2002
LOCUS BU383609 Dictyostelium discoideum cDNA library, CF Dictyostelium
DEFINITION dictoideum cDNA clone ddc5j18 3', mRNA sequence.
ACCESSION BU383609
VERSION BU383609.1 GI:19292993
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the culmination stage
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source Location/Qualifiers
1..612
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddc5j18"
/sex="mat A"
/dev_stage="Culmination stage"
/clone_id="Dictyostelium discoideum cDNA library, CF"
BASE COUNT 142 a 90 c 131 g 243 t 6 others
ORIGIN

Alignment Scores:
Pred. No.: 0.000926 Length: 612
Score: 168.50 Matches: 40
Percent Similarity: 48.96% Conservative: 7
Best Local Similarity: 41.67% Mismatches: 47
Query Match: 16.47% Indels: 2
DB: 12 Gaps: 1

US-10-081-935-4 (1-194) x BU383609 (1-612)

Oy 89 ProtleasPileProProValAspThrThralaValProGluProSerGluThrAlaGlu 108
Db 602 CCAACAGAGACTCCAACTGAGACTGAGACTGAGACTGAGACTGAGACTGAGACT 543
Oy 109 ProthralaGlu-ProthrgluGluProthralaGluProthralaGluProthralaGlu 128
Db 542 CCAACAGAGACTCCAACTGAGACTGAGACTGAGACTGAGACTGAGACTGAGACT 483
Oy 128 uProthriagluProthrgluGluProthralaValProthrglyThrGlyGlyVal 148
Db 482 TCCAAAGAAACTCCAAAGAAACTCCAAAGAAACTCCAAAGAAACTCCAAAGAAAC 423

Oy 148 lProthrglyThrGlySerPheThrValThrGlyArgProthralaSerThProAlaGlu 168
Db 422 TCCAAACGAGACTCCAACTGAAACTGTAATCTTACCAACCCCAACAGTA---ACACCAACTGA 366
Oy 168 uPheProGlyAlaGlySerAenValArgAlaSerValGlyGlyLe 183
Db 365 AACTCATCAAGTGAGAACTTTATCAATCATATAAAGTGATTA 320

RESULT 20
CNS05RNM
LOCUS CNS05RNM 1038 bp DNA linear GSS 01-SEP-2000
DEFINITION Tetradodon nigroviridis genome survey sequence T7 end of clone
014N06 of library C from Tetradodon nigroviridis, genomic survey
sequence.
ACCESSION AL350744
VERSION AL350744.1 GI:8244514
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradodon nigroviridis
ORGANISM Tetradodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetradodon.
1
Roest Crolius, H., Jailon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Broctier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
Estimate of human gene number provided by genome-wide analysis
using Tetradodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
20296633
10835645

REFERENCE
AUTHORS
TITLE
JOURNAL MEDLINE
PUBMED 20296633
10835645

2
Roest Crolius, H., Jailon, O., Dasilva, C., Ozouf-Costaz, C.,
Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Saurin, W., Bernot, A. and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)

JOURNAL MEDLINE
PUBMED 10899143
10899143

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetradodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetradodon.

FEATURES
source Location/Qualifiers
1..1038
/organism="Tetradodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="014N06"
/clone_id="C"
/note="Genoscope sequence ID : C07C014DG03C1-end : T7"
BASE COUNT 295 a 253 c 248 g 226 t 16 others
ORIGIN

Alignment Scores:
Pred. No.: 0.00181 Length: 1038
Score: 168.00 Matches: 32
Percent Similarity: 51.16% Conservative: 12
Best Local Similarity: 37.21% Mismatches: 42
Query Match: 16.42% Indels: 0
DB: 29 Gaps: 0

US-10-081-935-4 (1-194) x CNS05RNM (1-1038)

Oy 86 AlaGlyValProtleasPileProProValAspThrThralaValProGluProSerGlu 105


```

Db      561 GGTCTCTGCGCTCAAGGCTCTCGTCC-----CTTCTCATGGCTGCTTTC 608
RESULT 24
LOCUS   BU065834                      634 bp    mRNA    linear    EST 26-AUG-2002
DEFINITION   Fgr_8_K04_T3 Nitrogen-starved mycelia Gibberella zeae cDNA, mRNA
ACCESSION   BU065834
VERSION     BU065834.1  GI:22506123
KEYWORDS    EST.
SOURCE      Gibberella zeae
ORGANISM    Gibberella zeae
REFERENCE   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS     Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
TITLE       1 (bases 1 to 634)
            Trail,F., Xu,J.-R., San Miguel,P., Halgen,R.G. and Kistler,H.C.
            Analysis of expressed sequence tags from Gibberella zeae (anamorph
            Fusarium graminearum)
JOURNAL     Fungal Genet. Biol. 38 (2), 187-197 (2003)
MEDLINE     22508120
PUBMED      12620255
COMMENT     Contact: Frances Trail
            Department of Plant Biology
            Michigan State University
            East Lansing, MI 48824, USA
            Tel: 517 432 2939
            Fax: 517 353 1926
            Email: trail@msu.edu
            Plate: 8 row: K column: 04.
FEATURES
    source          1..634
                     location/Qualifiers
                     /organism="Gibberella zeae"
                     /mol_type="mRNA"
                     /strain="NRRL 31084"
                     /db_xref="taxon:5118"
                     /clone_lib="Nitrogen-starved mycelia"
                     /note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      128 a      231 c      154 g      121 t
ORIGIN
Alignment Scores:
Pred. No.:      0.00126      Length:      634
Score:          167.00      Matches:      53
Percent Similarity: 42.93%      Conservative: 32
Best Local Similarity: 26.77%      Mismatches: 89
Query Match:    16.32%      Indels:      24
DB:             13      Gaps:      8
US-10-081-935-4 (1-194) x BU065834 (1-634)
Oy      1 MetGlnPheSerHisAlaLeuIleAlaValAlaIaGlyLeuAlaSerAlaGlnLeu 20
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      27 ATGAAGTACTCGGCTGCTTGTGCTCTCGCGCTGTGCGCTCAGGCTCAGTCTCT 86
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Oy      21 ProAspIleProProCysAlaLeuAsnCysPheValGluAlaLeuGlyAAsnAspGlyCys 40
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      87 GCCACGTCGCCCAAGTGGCCATTCCTTGCCTGCACAAAGCCATGCCACGACGAC 146
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Oy      41 ThrArgLeuThrAspPheLysCysHisCysSerLys---ProGluLeuProGlyGlnIle 59
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      147 TGTGACAGACCGACTCGCGCTGTCTGTCGAAGGGCTTCAGCGCGTGTCCGATCAAGGCT 206
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Oy      60 ThrProCysValGluGluAlaCysProLeuAsnAlaArgIleSerValSerAsnIleVal 79
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      207 ACCCTTGTGTGATTGACGAGTGTGATCCGAC-----GTTGCCATCAACGAGTGC-- 257
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Oy      80 ValAspGlnCysSerLysAlaGlyValProIleAspIleProProValAspThrThrAla 99
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      258 -----CTTCCCGCTACCGACGAACTCTCC 281
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Oy      100 AlaProGluProSerGluThr---AlaGluProThrAlaGluProThrGluGluProThr 118
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

Db      282 AAGAACCCCGCCCAAGATGTGAGCCCAAGTCCACCGCCGAG---GAGAGAACCCACC 338
Oy      119 AlaGluProThrAlaGluProThr-----AlaGluProThrHisGluProThrGlu 135
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      339 ACCACCGCGCGCTGCACCTTCCTGCTCTGTACACACGACGCGGAGCTTTAG 398
Oy      136 GluProThrAlaValProThrGlyThrGlyGlyValProThrGlyThrGlySerPhe 155
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      399 ACCACCGCGCTGCACACCACTGCTCCCTCCATCTCCACCAACCGCTCTAGAG 458
Oy      156 ThrValThrGlyArgProThrAlaSerThrProAlaGluPheProGlyAlaGlySerAsn 175
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      459 CCGGACACGACACCGCGCGCGCTGCACCCCTACCAAGGATGCC---GAGCAGCCAAC 515
Oy      176 ValArgAlaSerValGlyGlyIleAlaAlaLeuLeuGlyLeuAlaIleVal 193
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      516 GGTCTCTGCGCTCAAGGCTCTCGTCC-----CTTCCCATGGCTGCTTTC 563
RESULT 25
LOCUS   BU064409                      657 bp    mRNA    linear    EST 26-AUG-2002
DEFINITION   Fgr_4_M16_T3 Nitrogen-starved mycelia Gibberella zeae cDNA, mRNA
ACCESSION   BU064409
VERSION     BU064409
KEYWORDS    EST.
SOURCE      Gibberella zeae
ORGANISM    Gibberella zeae
REFERENCE   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS     Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
TITLE       1 (bases 1 to 657)
            Trail,F., Xu,J.-R., San Miguel,P., Halgen,R.G. and Kistler,H.C.
            Analysis of expressed sequence tags from Gibberella zeae (anamorph
            Fusarium graminearum)
JOURNAL     Fungal Genet. Biol. 38 (2), 187-197 (2003)
MEDLINE     22508120
PUBMED      12620255
COMMENT     Contact: Frances Trail
            Department of Plant Biology
            Michigan State University
            East Lansing, MI 48824, USA
            Tel: 517 432 2939
            Fax: 517 353 1926
            Email: trail@msu.edu
            Plate: 4 row: M column: 16.
FEATURES
    source          1..657
                     location/Qualifiers
                     /organism="Gibberella zeae"
                     /mol_type="mRNA"
                     /strain="NRRL 31084"
                     /db_xref="taxon:5118"
                     /clone_lib="Nitrogen-starved mycelia"
                     /note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      137 a      239 c      155 g      126 t
ORIGIN
Alignment Scores:
Pred. No.:      0.00131      Length:      657
Score:          167.00      Matches:      53
Percent Similarity: 42.93%      Conservative: 32
Best Local Similarity: 26.77%      Mismatches: 89
Query Match:    16.32%      Indels:      24
DB:             13      Gaps:      8
US-10-081-935-4 (1-194) x BU064409 (1-657)
Oy      1 MetGlnPheSerHisAlaLeuIleAlaValAlaIaGlyLeuAlaSerAlaGlnLeu 20
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      72 ATGAAGTACTCGGCTGCTTGTGCTCTCGCGCTGTGCGCTCAGGCTCAGTCTCT 131
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Oy      21 ProAspIleProProCysAlaLeuAsnCysPheValGluAlaLeuGlyAAsnAspGlyCys 40
         |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      132 GCCACGTCGCCCAAGTGGCCATTCCTTGCCTGCACAAAGGCCATGCCGACGAC 191

```



```

Oy      41 ThrAgluThrAlaAspPheLeuCyHisIAsySerLys----ProGluLeuProGlyGlnIle 59
           |||||
Db      192 TGTGCAAGACCGGAGCTCGCTGCTGTCTGTCCAAAGGGCTTCAGCGCTCCGATCCAAAGGCT 251
           |||||
Oy      60 ThrProCyValGluGluAlaCyAProLeuAspAlaArgIleSerValSerAsnIleVal 79
           |||||
Db      252 ACCCTGTGTGTCATGATGACGAGTGTGATCCGAC-----GTTCGATCAACGAGGTCC-- 302
           |||||
Oy      80 ValAaGlnCysSerLysAlaGlyValProIleAspIleProProValAspThrAla 99
           |||||
Db      303 -----CTTCCGGCTAACGGAAAGCTCTGC 326
           |||||
Oy      100 AlaProGluProSerGluThr--AlaGluProThrAlaGluProThGlnIleProTh 118
           |||||
Db      327 AAGAACCCCCCAAGAGAGTGTGAGGCCAAAGTCAACGCCGAG--GAGAGAAAGCCAC 383
           |||||
Oy      119 AlaGluProThrAlaGluProThr-----AlaGluProThnIleGluProThGlu 135
           |||||
Db      384 ACCACCGCGCGCGCCACCTCCACCTTGTCGTGTGCACCCAGCCAGCGCGAGAGTTGTAG 443
           |||||
Oy      136 GluProThrAlaValProThGlyThrGlyGlyValProThGlyThrGlySerPhe 155
           |||||
Db      444 ACCACCGCGCGCTGCCACCACTGTGCTGCCATCATTCGCCACACCGCTGTCCAGAG 503
           |||||
Oy      156 ThrValThnGlyArgProThrAlaSerThrProAlaGluPheProGlyAlaGlySerAsn 175
           |||||
Db      504 CCGCGCACACACACCCCGCGCTGCCACACCTTACCAAGAGGCC--GAGCAGGCCAAC 560
           |||||
Oy      176 ValArgAlaSerValGlyGlyIleAlaAlaLeuLeuGlyLeuAlaAlaArg 193
           |||||
Db      561 GGTGCTGTGGCTTCAAGAGTGTCTGCTGCG-----CTGCGCATGAGTGGTTTC 608
           |||||

```

RESULT 26	
LOCUS	B060530
DEFINITION	B060530 664 bp mRNA linear EST 26-AUG-2002
ACCESSION	Fg-C_1_E06_T3 Carbon-starved mycelia Gibberella zeae cDNA, mRNA sequence.
VERSION	B060530.1
KEYWORDS	GI:22500819
SOURCE	EST.
ORGANISM	Gibberella zeae
REFERENCE	Gibberella zeae
AUTHORS	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
TITLE	1 (bases 1 to 664)
JOURNAL	Trail, F., Xu, J.-R., San Miguel, P., Halgren, R.G. and Kistler, H.C.
MEDLINE	Analysis of expressed sequence tags from Gibberella zeae (anamorph
PUBMED	Fusarium graminearum)
COMMENT	Fungal Genet. Biol. 38 (2), 187-197 (2003)
	22508120
	12620255
	Contact: Frances Trail
	Department of Plant Biology
	Michigan State University
	East Lansing, MI 48824, USA
	Tel: 517 432 2939
	Fax: 517 353 1926
	Email: trail@msu.edu
	Plate: 1 row: E column: 06.
FEATURES	location/Qualifiers
source	1..664
	/organism="Gibberella zeae"
	/mol_type="mRNA"
	/strain="NRRL 31084"
	/db_xref="taxon:5518"
	/clone_id="Carbon-starved mycelia"
	/note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT	139 a 241 c 155 g 129 t
ORIGIN	
Alignment Scores:	0.00132 Length: 664

Score:	167.00	Matches:	53
Percent Similarity:	42.93%	Conservative:	12
Best Local Similarity:	26.77%	Mismatches:	89
Query Match:	16.32%	Indels:	24
DB:	13	Gaps:	8

```

Qy      1 MetcIlnheerhlsalaleuillalaleuvalAlaAlaglyLeuAlaserIaglnleu 20
Db      80 ATGAAGTACTCCGCTGCTTTCGTTGCTCTCCGCGCTTCCGCGCTCAGGCTAGTCTTT 139
Qy      21 ProAhpIleProProCyAlaLeuasnCyAspHeValGuaAlaLeuGlyAsnAhpIlyCyv 40
Db      140 GCGACGCTCCCAAGTGGCGGCATTCTTGCTCGACAGGCATCGCCAGCGACCAAC 199
Qy      41 ThirAglLeuThrAhpHeIyCyAspIleCySerIys---ProGlyLeuProGlycIlnIle 59
Db      200 TGTGCAAGACCGGACTCGCGCTGTGTGCAAGGGCTTCACGCGCTGCATCCAAAGCT 259
Qy      60 ThrProCyAlaGluGlnAlaCyvProLeuasnAlaArgIleSerValSerAsnIleVal 79
Db      260 ACCCTTGTGTCTATTGACGAGTGTGTGAACGAC---GTTGCCATTAAACAGGCT--- 310
Qy      80 ValAspIleCySerIyAlaGlyValProIleAhpIleProProValAspThrThrala 99
Db      311 -----CTTCCGCGTACCGAGAACTTCC 334
Qy      100 AlaProGluProSerGlyThr--AlaGluProThralaGluProThrGluGluProThr 118
Db      335 AAGAACCCCCCAAGAGGTGTGAGCCCAAGTCCACCGCGGAG--GAGGAGAAACCCACC 391
Qy      119 AlaGluProThralaGluProThr-----AlaGluProThrHisGluProThrGlu 135
Db      392 ACCACGCGCGGTGACCACTCCACTTGGTCTGTGTACACACAGCGCCGAGGTTGTGAG 451
Qy      136 GluProThralaValProThrGlyThrGlyGlyValProThrGlyThrGlySerPhe 155
Db      452 ACCACGCGCGGTGACCAACACATGTGCGTCCCATTCATCCACACCGCTGTGAGGAG 511
Qy      156 ThrValThrGlyArgProThrAlaSerThrProAlaGluPheProGlyValGlySerAsn 175
Db      512 CCGCGCACCGAACCCCGCGCGTGCACACCCCTTACCAAGGTGCC--GACACAGCCAAC 568
Qy      176 ValArgAlaSerValGlyGlyIleAlaAlaAlaLeuasnGlyLeuAlaAlaIleTy 193
Db      569 GGTGCTGCTGCTCTCAAGGTCTCGTGTCC---CTCGGCATGTGCTGCTTTT 616

```

RESULT	27
LOCUS	BU065008
DEFINITION	BU065008 696 bp mRNA linear EST 26-AUG-2002 Fgr_6_GG7_T3 Nitrogen-starved mycelia Gibberella zeae cDNA, mRNA sequence.
ACCESSION	BU065008
VERSION	BU065008.1 GI:22505297
KEYWORDS	EST.
SOURCE	Gibberella zeae
ORGANISM	Gibberella zeae
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypotracheales; Nectriaceae; Gibberella. 1 (bases 1 to 696)
AUTHORS	Traill,F., Xu,J.-R., San Miguel,P., Halgren,R.G. and Kistler,H.C. Analysis of expressed sequence tags from Gibberella zeae (anamorph Fusarium graminearum)
JOURNAL	Fungal Genet. Biol. 38 (2), 187-197 (2003)
MEDLINE	22508120
PUBMED	12620255
COMMENT	Contact: Frances Traill Department of Plant Biology Michigan State University East Lansing, MI 48824, USA Tel: 517 432 2939 Fax: 517 353 1926

Email: trail@msu.edu
 Plate: 6 row: G column: 07.
 Location/Qualifiers
 1. .696
 /organism="Gibberella zeae"
 /mol_type="mRNA"
 /db_xref="NRRL 31084"
 /clone_idb="Nitrogen-starved mycelia"
 /note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 151 a 248 c 162 g 135 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.00139 Length: 696
 Score: 167.00 Matches: 53
 Percent Similarity: 42.93% Conservative: 32
 Best Local Similarity: 26.77% Mismatches: 89
 Query Match: 16.32% Indels: 24
 DB: 13 Gaps: 8

US-10-081-935-4 (1-194) x BU065008 (1-696)

QY 1 MetGlnPheSerHISAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu 20
 DB 80 ATGAGTACTCTCGTCTGCTTCTGCTCTGCGCGCTGTTGCCGCTCAGGCTCAGTCTCTT 139
 QY 21 ProAspIleProProCysAlaLeuAlaLeuAlaLeuGlyAlaAspIleCys 40
 DB 140 GCCGACGTCCCAAGTGGCCATCTCTGCTGCAAGAGCCATCGCCAGCCAGCCAGCCAGC 199
 QY 41 ThrArgLeuThrAspPheLeuCysHisCysSerLeu---ProGluLeuProGlyGlnIle 59
 DB 200 TGTGACAAAGACCGACCTCGCTGCTGTCGCAAGGCGCTTCAGCGCTCGATCCAGGCT 259
 QY 60 ThrProCysValGluGluAlaCysProLeuAspAlaArgIleSerValSerAlaIleVal 79
 DB 260 ACCTTGTGTCTATTGACGAGTGTGTACCCGAC-----GTGCCATCAACAGGCTC--- 310
 QY 80 ValAspGlnCysSerLeuAlaGlyValProIleAspIleProProValAspThrThrAla 99
 DB 311 -----CTTCCGCTAACGAGAACCTCTGC 334
 QY 100 AlaProGluProSerGluThr---AlaGluProThrAlaGluProThrGluGluProThr 118
 DB 335 AAGAACCCGCCCAAGAGCTGAGGCCAAGTCACCGCCGAG--GAGGAGAAGCCCAAC 391
 QY 119 AlaGluProThrAlaGluProThr-----AlaGluProThrHisGluProThrGlu 135
 DB 392 ACCACCGCGCGTCCACCTCCGCTTGTGCTGTCACACAGCCGCGAGGTTGTGAG 451
 QY 136 GluProThrAlaValProThrGlyThrGlyGlyValProThrGlyThrGlySerPhe 155
 DB 452 ACCACCGCGCGTCCACCTCCGCTGCTCCATCATTCACACACCCGCTGAGAGAG 511
 QY 156 ThrValThrGlyArgProThrAlaSerThrProAlaGluPheProGlyAlaIleSerAla 175
 DB 512 CCGCGCACGAGACCCCGCGCTGCCACCTTACCAAGGATCC--GAGCAGGCCCAAC 568
 QY 176 ValArgAlaSerValGlyGlyIleAlaAlaLeuLeuGlyLeuAlaIleVal 193
 DB 569 GGTGCTGCTGCGCTCAAGGCTCTCGTGCC-----CTGCCATGCTCTTCTT 616

RESULT 28
 BU066113 699 bp mRNA linear EST 26-AUG-2002
 LOCUS Fgr 9 G08_T3 Nitrogen-starved mycelia Gibberella zeae cDNA, mRNA
 DEFINITION sequence.
 ACCESSION BU066113
 VERSION BU066113.1 GI:22506402
 KEYWORDS EST.
 SOURCE Gibberella zeae
 ORGANISM Gibberella zeae

REFERENCE
 AUTHORS Trail, F., Xu, J., R., San Miguel, P., Halgren, R.G. and Kistler, H.C.
 TITLE Analysis of expressed sequence tags from *Gibberella zeae* (anamorph *Fusarium graminearum*)
 JOURNAL Fungal Genet. Biol. 38 (2), 187-197 (2003)
 MEDLINE 22508120
 PUBMED 12620255

COMMENT

Contact: Frances Trail
 Department of Plant Biology
 Michigan State University
 East Lansing, MI 48824, USA
 Tel: 517 432 2939
 Fax: 517 353 1926
 Email: trail@msu.edu
 Plate: 9 row: G column: 08.
 Location/Qualifiers
 1. .699
 /organism="Gibberella zeae"
 /mol_type="mRNA"
 /strain="NRRL 31084"
 /db_xref="Taxon:5518"
 /clone_idb="Nitrogen-starved mycelia"
 /note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"

FEATURES

source

BASE COUNT 152 a 249 c 162 g 136 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.0014 Length: 699
 Score: 167.00 Matches: 53
 Percent Similarity: 42.93% Conservative: 32
 Best Local Similarity: 26.77% Mismatches: 89
 Query Match: 16.32% Indels: 24
 DB: 13 Gaps: 8

US-10-081-935-4 (1-194) x BU066113 (1-699)

QY 1 MetGlnPheSerHISAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu 20
 DB 82 ATGAGTACTCTCGTCTGCTTCTGCTCTGCGCGCTGTTGCCGCTCAGGCTCAGTCTCTT 141
 QY 21 ProAspIleProProCysAlaLeuAlaLeuAlaLeuGlyAlaAspIleCys 40
 DB 142 GCCGACGTCCCAAGTGGCCATCTCTGCTGCAAGGCGCTGCCAGCCAGCCAGCCAGC 201
 QY 41 ThrArgLeuThrAspPheLeuCysHisCysSerLeu---ProGluLeuProGlyGlnIle 59
 DB 202 TGTGACAAAGACCGACCTCGCTGCTGTCGCAAGGCGCTTCAGCGCTCGATCCAGGCT 261
 QY 60 ThrProCysValGluGluAlaCysProLeuAspAlaArgIleSerValSerAlaIleVal 79
 DB 262 ACCTTGTGTCTATTGACGAGTGTGTACCCGAC-----GTGCCATCAACAGGCTC--- 312
 QY 80 ValAspGlnCysSerLeuAlaGlyValProIleAspIleProProValAspThrThrAla 99
 DB 313 -----CTTCCGCTAACGAGAACCTCTGC 336
 QY 100 AlaProGluProSerGluThr---AlaGluProThrAlaGluProThrGluGluProThr 118
 DB 337 AAGAACCCGCCCAAGAGCTGAGGCCAAGTCACCGCCGAG--GAGGAGAAGCCCAAC 393
 QY 119 AlaGluProThrAlaGluProThr-----AlaGluProThrHisGluProThrGlu 135
 DB 394 ACCACCGCGCGTCCACCTCCGCTTGTGCTGTCACACAGCCGCGAGGTTGTGAG 453
 QY 136 GluProThrAlaValProThrGlyThrGlyGlyValProThrGlyThrGlySerPhe 155
 DB 454 ACCACCGCGCGTCCACCACTGCTGCTCCATCATTCACACACCGCTGAGAGAG 513
 QY 156 ThrValThrGlyArgProThrAlaSerThrProAlaGluPheProGlyAlaIleSerAla 175
 DB 514 CCGCGCACGAGACCCCGCGCTGCCACCTTACCAAGGATCC--GAGCAGGCCCAAC 570

```

Qy      176 ValArgAlaSerValGlyGlyIleAlaAlaLeuLeuGlyLeuAlaAlaTyr 193
Db      571 GGTGCTCTGCGCTCAAGGCTCTGTCGCC-----CTCCGCAATGCGTCTTTC 618

RESULT 29
LOCUS   BU064737              718 bp    mRNA           linear    EST 26-AUG-2002
DEFINITION fgr_5_K14_T3 Nitrogen-starved mycelia Gibberella zeae cDNA, mRNA
ACCESSION BU064737
VERSION   BU064737.1    GI:22505026
KEYWORDS  EST.
SOURCE    Gibberella zeae
ORGANISM  Gibberella zeae
          Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
          Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE 1 Trail,F., Xu,J.-R., San Miguel,P., Halgren,R.G. and Kistler,H.C.
          Analysis of expressed sequence tags from Gibberella zeae (anamorph
          Fusarium graminearum)
JOURNAL   Fungal Genet. Biol. 38 (2), 187-197 (2003)
MEDLINE   12620255
PUBMED    12620255
COMMENT   Contact: Frances Trail
          Department of Plant Biology
          Michigan State University
          East Lansing, MI 48824, USA
          Tel: 517 432 2939
          Fax: 517 353 1926
          Email: trail@msu.edu
          Plate: 5 row: K column: 14.
FEATURES
  source          1..718
                  Location/Qualifiers
                  /organism="Gibberella zeae"
                  /mol_type="mRNA"
                  /strain="NRRL 31084"
                  /db_xref="taxon:5518"
                  /clone_lib="Nitrogen-starved mycelia"
                  /note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      157 a      253 c      168 g      140 t
ORIGIN
Alignment Scores:
Pred. No.:      0.00144      Length:      718
Score:          167.00      Matches:      53
Percent Similarity: 42.93%      Conservative: 32
Best Local Similarity: 26.77%      Mismatches: 89
Query Match:    16.32%      Indels:      24
DB:             13      Gaps:      8
US-10-081-935-4 (1-194) x BU064737 (1-718)
Qy      1 MetGlnPheSerIleAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu 20
Db      80 ATGAAGACTCCCGCGCTTTCGTCGTCCTGCGCGCTGTCGCCGCTCAAGCTCACTCTT 139
Qy      21 ProAspIleProProCysAlaLeuAenCysPheValAlaAlaLeuGlyAsnAspGlyCys 40
Db      140 GCCAGCTGCCCAAGTCGCCATTCCTGCTCCACAGAGCATCGCCAGCGAGACGAC 199
Qy      41 ThrArgLeuThrAspPheLysCysHisCysSerLys---ProGluLeuProGlyGlnIle 59
Db      200 TGTGACAAAGCCACCTCGCGCTGTGTCTGCAAGGGCTTCACCGCTGTCATCCAAAGCT 259
Qy      60 ThrProCysValGluGluAlaCysProLeuAspAlaArgIleSerValSerAsnIleVal 79
Db      260 ACCTCTGTGTCAATTGACGAGTGTGTAACGAC-----GTTCGCAATCAACGAGTC--- 310
Qy      80 ValAspGlnCysSerLysAlaGlyValProIleAspIleProProValAspThrThrAla 99
Db      311 -----CTTCCCGCTACCGAAGACTCTGTC 334

```

```

Qy      100 AlaProGluProSerGluThr---AlaGluProThrAlaGluProThrGluGluProThr 118
Db      335 AAGAACCCCCCAAGAGGTCTGAGGCCAAGTCCACCCCGCAG--GACGAGAACCCACCC 391
Qy      119 AlaGluProThrAlaGluProThr-----AlaGluProThrHisGluProThrGlu 135
Db      392 ACCACGCCCGCTGACCTCCACTTGTGTCCTTTCACACACAGCCGCGGAGTTCTTAG 451
Qy      136 GluProThrAlaValAlaProThrGlyThrGlyGlyValAlaProThrGlyThrGlySerPhe 155
Db      452 ACCACCCCGCTGCGACACCATCTGCTCCATCTCCACACCCGCTGTGAGAG 511
Qy      156 ThrValThrGlyArgProThrAlaSerThrProAlaGluPheProGlyAlaGlySerAsn 175
Db      512 CCCCCCAGACAGACCCCGCGCTGCCACCCCTTACCAAGGCTCC--GACAGGCCAAC 568
Qy      176 ValArgAlaSerValGlyGlyIleAlaAlaLeuLeuGlyLeuAlaAlaTyr 193
Db      569 GGTGCTCTGCGCTCAAGGCTCTGTCGCC-----CTCCGCAATGCGTCTTTC 616

RESULT 30
LOCUS   BU064923              726 bp    mRNA           linear    EST 26-AUG-2002
DEFINITION fgr_6_C13_T3 Nitrogen-starved mycelia Gibberella zeae cDNA, mRNA
ACCESSION BU064923
VERSION   BU064923.1    GI:22505212
KEYWORDS  EST.
SOURCE    Gibberella zeae
ORGANISM  Gibberella zeae
          Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
          Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE 1 Trail,F., Xu,J.-R., San Miguel,P., Halgren,R.G. and Kistler,H.C.
          Analysis of expressed sequence tags from Gibberella zeae (anamorph
          Fusarium graminearum)
JOURNAL   Fungal Genet. Biol. 38 (2), 187-197 (2003)
MEDLINE   12620255
PUBMED    12620255
COMMENT   Contact: Frances Trail
          Department of Plant Biology
          Michigan State University
          East Lansing, MI 48824, USA
          Tel: 517 432 2939
          Fax: 517 353 1926
          Email: trail@msu.edu
          Plate: 6 row: C column: 13.
FEATURES
  source          1..726
                  Location/Qualifiers
                  /organism="Gibberella zeae"
                  /mol_type="mRNA"
                  /strain="NRRL 31084"
                  /db_xref="taxon:5518"
                  /clone_lib="Nitrogen-starved mycelia"
                  /note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      159 a      256 c      168 g      143 t
ORIGIN
Alignment Scores:
Pred. No.:      0.00146      Length:      726
Score:          167.00      Matches:      53
Percent Similarity: 42.93%      Conservative: 32
Best Local Similarity: 26.77%      Mismatches: 89
Query Match:    16.32%      Indels:      24
DB:             13      Gaps:      8
US-10-081-935-4 (1-194) x BU064923 (1-726)
Qy      1 MetGlnPheSerIleAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu 20
Db      88 ATGAAGACTCCCGCGCTTTCGTCGTCCTGCGCGCTGTCGCCGCTCAAGCTCACTCTT 147
Qy      21 ProAspIleProProCysAlaLeuAenCysPheValAlaAlaLeuGlyAsnAspGlyCys 40

```

```

Db      148 GCCGACGTCCTCCAGTGCCTGCTTCTGCTGACACAGCCATCCGACCGAGACCAGC 207
Qy      41 ThrArgLeuThrAspPheLysCysHisCysSerLys---ProGluLeuProGlyGlnIle 59
Db      208 TGTGACAGAGCCGACCTCGCTGTGTGACAGGCGCTTACAGCGCTGTCATCCATCAAGGCT 267
Qy      60 ThrProCysValGluGluAlaCysProLeuAspAlaArgIleSerValSerAsnIleVal 79
Db      268 ACCTCTGTGTCTCATTTGACGAGTGTGTGACCGAC-----GTTCGCAATCAACGAGGTC--- 318
Qy      80 ValAspGlnCysSerLysValGlyValProIleAspIleProProValAspThrThrAla 99
Db      319 -----CTTCCCGCTACCGAGAACCTCTGC 342
Qy      100 AlAProGluProSerGluThr---AlaGluProThrAlaGluProThrGluGluProThr 118
Db      343 AAGAACCCCCCAAGAGGTGTGAGGCCAAGTCCACCCCGAG---GAGGAGAACCCACAC 399
Qy      119 AlaGluProThrAlaGluProThr-----AlaGluProThrHisGluProThrGlu 135
Db      400 ACCACCCCGCTGCGACCTCCACCTGTGCTGTGTGCACACACGCGCGAGGTTGTAG 459
Qy      136 GluProThrAlaValProThrGlyThrGlyGlyValProThrGlyThrGlySerPhe 155
Db      460 ACCACCCCGCTGCGACCTCCACCTGTGCTGTGTGCACACACGCGCGAGGTTGTAG 519
Qy      156 ThrValThrGlyArgProThrAlaSerThrProAlaGluPheProGlyAlaGlySerAsn 175
Db      520 CCGGCGACGACACCCCGCGGTGCGACCCCTTACCAAGGCTCC---GAGCAGGCCAAC 576
Qy      176 ValArgAlaSerValGlyGlyIleAlaAlaAlaLeuLeuGlyLeuAlaIleTyr 193
Db      577 GTGCTCTGCTGCGCTCAAGGCTGTGCTGCC-----CTCCGCAATGCTGCTTTC 624

RESULT 31
CD458964      841 bp      mRNA      linear      EST 03-JUN-2003
DEFINITION   Fg08_01h09_A Fg08_AAFC_ECORC_Fusarium_graminearum_complex_substrate
Gibberella zeae cDNA clone Fg08_01h09, mRNA sequence.
ACCESSION   CD458964
VERSION     CD458964.1 GI:31373704
KEYWORDS    EST.
SOURCE      Gibberella zeae
ORGANISM    Gibberella zeae
REFERENCE   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS     Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
            Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J., Hattori,
            J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A., Sprott,D. and
            Tinker,N.A.
            A cDNA library prepared from Fusarium graminearum grown on a
            complex plant substrate
            Unpublished
            Contact: Watson, Robert.J.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
            CANADA
            Tel: (613) 759-1655
            Fax: (613) 759-1701
            Email: watsonrj@agr.gc.ca.

FEATURES
    source
        1..841
            /organism="Gibberella zeae"
            /mol_type="mRNA"
            /strain="DAOM 180378"
            /db_xref="taxon:5518"
            /clone="Fg08_01h09"
            /tissue_type="Mycelium"
            /dev_stage="Asexual"
            /lab_host="E. coli DH10B"
            /clone_id="Fg08_AAFC_ECORC_Fusarium_graminearum_complex_s
```

```

                                substrate"
                                /note="Vector: pBluescript II+; Site 1: EcoRI; Site 2:
                                XhoI; Fusarium graminearum grown on a complex plant_
                                substrate-- wheat leaves treated to remove most of the low
                                molecular weight, water-soluble components."
BASE COUNT      191 a      270 c      200 g      179 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      0.00172      Length:      841
Score:          167.00      Matches:      53
Percent Similarity: 42.93%      Conservative: 32
Best Local Similarity: 26.77%      Mismatches: 89
Query Match:    16.32%      Indels:      24
Db:             14      Gaps:      8

US-10-081-935-4 (1-194) x CD458964 (1-841)
Qy      1 MetGlnPheSerHisAlaLeuIleAlaLeuValAlaIleGlyLeuAlaSerAlaGlnLeu 20
Db      82 ATGAGTACTTCCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 141
Qy      21 ProAspIleProProCysAlaLeuAsnCysPheValGluAlaLeuGlyAsnAspGlyCys 40
Db      142 GCCGAGCTCCCAAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 201
Qy      41 ThrArgLeuThrAspPheLysCysHisCysSerLys---ProGluLeuProGlyGlnIle 59
Db      202 TGTGACAGAGCCGACCTCGCTGTGTGTGCAAGGCTTACAGCGCTGTCCATCCACAGCT 261
Qy      60 ThrProCysValGluGluAlaCysProLeuAspAlaArgIleSerValSerAsnIleVal 79
Db      262 ACCTCTGTGTCTTGTGAGAGTGTGACGAC-----GTTCGCAATCAACGAGTCT--- 312
Qy      80 ValAspGlnCysSerLysValGlyValProIleAspIleProProValAspThrThrAla 99
Db      313 -----CTTCCCGCTACCGAGAACCTCTGC 336
Qy      100 AlAProGluProSerGluThr---AlaGluProThrAlaGluProThrGluGluProThr 118
Db      337 AAGAACCCCCCAAGAGGTGTGAGGCCAAGTCCACCCCGAG---GAGGAGAACCCACAC 393
Qy      119 AlaGluProThrAlaGluProThr-----AlaGluProThrHisGluProThrGlu 135
Db      394 ACCACCCCGCTGCGACCTCCACCTGTGCTGTGTGCACACACGCGCGAGGTTGTAG 453
Qy      136 GluProThrAlaValProThrGlyThrGlyGlyValProThrGlyThrGlySerPhe 155
Db      454 ACCACCCCGCTGCGACCTCCACCTGTGCTGTGTGCACACACGCGCGAGGAG 513
Qy      156 ThrValThrGlyArgProThrAlaSerThrProAlaGluPheProGlyAlaGlySerAsn 175
Db      514 CCGGCGACGACACCCCGCGGTGCGACCCCTTACCAAGGCTCC---GAGCAGGCCAAC 570
Qy      176 ValArgAlaSerValGlyGlyIleAlaAlaAlaLeuLeuGlyLeuAlaIleTyr 193
Db      571 GTGCTCTGCTGCGCTCAAGGCTGTGCTGCC-----CTCCGCAATGCTGCTTTC 618

RESULT 32
CD458020      867 bp      mRNA      linear      EST 03-JUN-2003
LOCUS       Fg08_01h09_R Fg08_AAFC_ECORC_Fusarium_graminearum_complex_substrate
DEFINITION   Gibberella zeae cDNA clone Fg08_01h09, mRNA sequence.
ACCESSION   CD458020
VERSION     CD458020.1 GI:31372760
KEYWORDS    EST.
SOURCE      Gibberella zeae
ORGANISM    Gibberella zeae
REFERENCE   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS     Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
            Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J., Hattori,
            J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A., Sprott,D. and
```

Tinker, N.A.
A cDNA library prepared from *Fusarium graminearum* grown on a complex plant substrate

Unpublished
Contact: Watson, Robert J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA

TEL: (613) 759-1655
FAX: (613) 759-1701
Email: watsonrj@agr.gc.ca.
Location/Qualifiers

FEATURES
source
1..867
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="taxon:5518"
/clone="Fg08_01h09"
/tissue_type="Mycelium"
/dev_stage="Asexual"
/lab_host="E. coli DH10B"
/clone_lib="Fg08_AAF_C_ECORC_Fusarium_graminearum_complex_s
ubstrate"
/note="Vector: pBluescript II+; Site 1: EcoRI; Site 2:
XhoI; *Fusarium graminearum* grown on a complex plant
substrate-- wheat leaves treated to remove most of the low
molecular weight, water-soluble components."

BASE COUNT 204 a 273 c 207 g 183 t
ORIGIN

Alignment Scores:

Pred. No.: 0.00178 length: 867
Score: 167.00 Matches: 53
Percent Similarity: 42.93% Conservative: 32
Best Local Similarity: 26.77% Mismatches: 89
Query Match: 16.32% Indels: 24
DB: 14 Gaps: 8

US-10-081-935-4 (1-194) x CD458020 (1-867)

```

QY      1 MetGlnPheSerHisAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu 20
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB      82 ATGAGTACTCGCTCGCTTGTGCTGCGCGCTGTCGCGCTCAGCTAGTCTT 141
QY      21 ProAspIleProProCysAlaLeuAsnCysPheValGlnAlaLeuGlyAsnAspGlyCys 40
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB      142 GCGGAGTCCCAAGTCCGCTTCTGCGATTCGACAGCGCAGACGAGCC 201
QY      41 ThrArgLeuThrAspPheLysCysHisCysSerLys---ProGluLeuProGlyGlnIle 59
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB      202 TGTGACAAACCGACCTCGCTGTGTCTGCAAGGCTTACGCCCTGTCGATCCAAAGCT 261
QY      60 ThrProCysValGlnGlnAlaCysProLeuAspAlaArgIleSerValSerAsnIleVal 79
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB      262 ACCCTTGTGTCTATTGACGAGTGTGTGACGAC-----GTTCGATCAACGAGGTC--- 312
QY      80 ValAspGlnCysSerLysAlaGlyValProIleAspIleProProValAspThrThrAla 99
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB      313 -----CTTCCCTCCCTACCGAGAACTCTCTCG 316
QY      100 AlaProGluProSerGluThr---AlaGluProThrAlaGluProThrGluGluProThr 118
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB      337 AAGAACCCCGCCAGAGAGTGTGAGCGCAAGTCCACCGCCGAG---GAGAGAAAGCCACCC 393
QY      119 AlaGluProThrAlaGluProThr-----AlaGluProThrHisGluProThrGlu 135
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB      394 ACCACCGCGCGTCCGACCTTGTGTGTGTCACACGCGCCGAGGTTGTTGAG 453
QY      136 GluProThrAlaValaProThrGlyThrGlyGlyValaProThrGlyThrGlySerPhe 155
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB      454 ACCACCGCGCGTCCGACCGACCTGTGCTCCATTCATTCACACCGCGCTGTGAGAG 513

```

QY 156 ThrValThrGlyArgProThrAlaSerThrProAlaGluPheProGlyAlaGlySerAsn 175
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 514 CCGCCACGACGACCCCGCGCTGCGCCCTTACCAAGGCTCC---GACGAGCCCAAC 570
QY 176 ValArgAlaSerValGlyGlyIleAlaAlaLeuLeuGlyLeuAlaAlaTyr 193
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 571 GGTGCTGCTGGGCTCAAGGCTTCGTCGCC-----CTCCGATGCTGCTTTC 618

RESULT 33

BM686215 606 bp mRNA linear EST 06-MAY-2003
mgce013xp19f.b Magnaporthe grisea CS Uni-Zap XR Library Magnaporthe
grisea cDNA clone mgce013xp19 5', mRNA sequence.

ACCESSION BM686215.2 GI:30393206
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

AUTHORS
TITLE
Unpublished
On Mar 7, 2002 this sequence version replaced gi:19235897.
Contact: Ebbole DJ
Department of Plant Pathology & Microbiology
Texas A&M University
Peterson Bldg, MS2132, College Station, TX 77843-2132, USA
Tel: 979 845 4831
Fax: 979 845 6483
Email: d-ebbole@tamu.edu
Chromatogram file of this sequence is available, see contact person
Best hit (April. 22, 2003) gb|AA64312.1| MAC1 interacting
protein 1, AC11 [Magnaporthe grisea] . 179 3e-44
PCR Primers
FORWARD: T3 primer
BACKWARD: T7 primer
Plate: mgce013 row: P column: 19
Seq primer: T3.

JOURNAL

COMMENT
Location/Qualifiers

FEATURES

1..606
/organism="Magnaporthe grisea"
/mol_type="mRNA"
/strain="Guy11"
/db_xref="taxon:148105"
/clone="mgce013xp19"
/sex="Mati-2 hermaphrodite"
/cell_type="conidia"
/note="Vector: pBluescriptSK-; Site 1: EcoRI; Site 2: XhoI
; unidirectional cloning. EcoRI site has T3 primer--and
predominantly 5' reads. T7 primer on XhoI site of insert.
Conidial library. Point inoculation of Guy11 at center of
oatmeal agar plate. Conidia were harvested after two weeks
of growth. Sequences were processed by one of two methods.
Where a full-length alignment to the M. grisea genome
sequence was available, the EST sequence was trimmed
according to the alignment, otherwise sequence quality was
assessed using phred/nap version 991019 and trimmed
according to pnd files (0.05) and for vector segs."

BASE COUNT

115 a 200 c 171 g 120 t
ORIGIN

Alignment Scores:

Pred. No.: 0.00143 length: 606
Score: 166.00 Matches: 54
Percent Similarity: 42.94% Conservative: 22
Best Local Similarity: 30.51% Mismatches: 67
Query Match: 16.23% Indels: 34
DB: 12 Gaps: 9

COMMENT

Contact: Watson, Robert.J.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-Food Canada
 Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
 CANADA
 Tel: (613) 759-1655
 Fax: (613) 759-1701
 Email: watsonrj@agr.gc.ca

FEATURES

Location/Qualifiers
 1..822

BASE COUNT 184 a 267 c 195 g 172 t 4 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.002 Length: 822
 Score: 166.00 Matches: 54
 Percent Similarity: 43.22% Conservative: 32
 Best Local Similarity: 27.14% Mismatches: 88
 Query Match: 16.23% Indels: 25
 DB: 14 Gaps: 8

/organism="Gibberella zeae"
 /mol_type="mRNA"
 /strain="DAOM 180378"
 /db_xref="taxon:5518"
 /clone="Fg09_03113"
 /tissue_type="mycelium"
 /dev_stage="asexual"
 /lab_host="E. coli DH10B"
 /clone_id="Fg09_AAFC_ECORC_Fusarium_graminearum_simple_su
 bstrate"
 /note="Vector: pBluescript II; Site_1: EcoRI, Site_2:
 XhoI; Fusarium graminearum grown on a simple substrate--
 minimal media supplemented with amino acids."

US-10-081-935-4 (1-194) x CD460963 (1-822)

QY 1 MetGlnPheSerHisAlaLeuValAlaGlyLeuAlaSerAlaGlnLeu 20
 DB 82 ATGAAGACTCCGTCGCTTCGTCGCTCCGCCGCTTCGCCGCTCACTCTT 141
 QY 21 ProAspIleProProCysAlaLeuAsnGlyPheValGlnAlaLeuGlyAsnAspGlyCys 40
 DB 142 GCCGACGTCCCAAGTCGCGCATCTTCCTCGACAGCCATCGCCAGCAGACGAGC 201
 QY 41 ThrArgLeuThrAspPheLysCysHisCysSerLys---ProGluLeuProGlyGlnIle 59
 DB 202 TGTGACAAAGCCACCTCGCTGTGTGTGCAAGGCTTCAGCGCTGCATCCATCCAGGCT 261
 QY 60 ThrProCysValGlnGluAlaCysProLeuAspAlaArgIleSerValSerAsnIleVal 79
 DB 262 ACCTCTTGTCATTGACGAGTGTGTAACGAC-----GTTGCCATCAACGAGGTC--- 312
 QY 80 ValAspGlnCysSerLysAlaGlyValProIleAspIleProProValAspThrThrAla 99
 DB 313 -----CTTCCCGCTACCGAAGACTCTGTC 336
 QY 100 AlaProGluProSerGluThr---AlaGluProThrAlaGluProThrGluGluProThr 118
 DB 337 AAGAACCCCCCAAGAGTGTGAGGCCAAGTCCACGCGCAG--GAGGAGAACCCACCC 353
 QY 119 Ala-GluProThrAlaGluPro-----ThrAlaGluProThrHisGluProThrGln 135
 DB 394 ACCACCGCGCTGCACCTCCACCTTGTCGTCGTCACCAAGGCGCGAGGTCGTTGTA 453
 QY 135 uGluProThrAlaValProThrGlyThrGlyGlyValProThrGlyThrGlySerPh 155
 DB 454 GACCAACCGCGCTGCACCACTGTGCTCCATCATTCACCAACCGCTCTGAGGA 513
 QY 155 eThrValThrGlyArgProThrAlaSerThrProAlaGluPhProGlyValAlaGlySerPh 175
 DB 514 GCCCGGACCAAGACCCCGCGCTGCACCTTACCAAGGCTCC---GAGCAGGCGCAA 570
 QY 175 nValArgAlaSerValGlyGlyIleAlaAlaLeuLeuGlyLeuAlaAlaIleArg 193

DB 571 CGGTCGTCGCGCTCAAGGTCCTCGCTGCC-----CTGCCATGCTGCTTTC 619

RESULT 36

BU064419 611 bp mRNA linear EST 26-AUG-2002
 LOCUS BU064419
 DEFINITION Fg4_N02_T3 Nitrogen-starved mycelia Gibberella zeae cDNA, mRNA
 sequence.

ACCESSION BU064419.1 GI:22504708
 VERSION BU064419.1
 KEYWORDS EST.

SOURCE

ORGANISM Gibberella zeae
 Gibberella zeae
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.

REFERENCE Trail,J., Xu,J.-R., San Miguel,P., Halgren,R.G. and Kistler,H.C.
 1 (bases 1 to 611)
 TITLE Analysis of expressed sequence tags from Gibberella zeae (anamorph

Fusarium graminearum)
 Fungal Genet. Biol. 38 (2), 187-197 (2003)

JOURNAL MEDLINE
 PUBMED 12620255

COMMENT Contact: Frances Trail
 Department of Plant Biology
 Michigan State University
 East Lansing, MI 48824, USA

Tel: 517 432 2939
 Fax: 517 353 1926
 Email: trail@msu.edu

Plate: 4 row: N column: 02.
 Location/Qualifiers

FEATURES

source

1..611
 /organism="Gibberella zeae"
 /mol_type="mRNA"
 /strain="NRRL 31084"
 /db_xref="taxon:5518"
 /clone_id="Nitrogen-starved mycelia"
 /note="Vector: Uni-ZapII; Site_1: EcoRI, Site_2: XhoI"

BASE COUNT 128 a 229 c 143 g 111 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.00188 Length: 611
 Score: 164.50 Matches: 52
 Percent Similarity: 40.61% Conservative: 28
 Best Local Similarity: 26.40% Mismatches: 88
 Query Match: 16.08% Indels: 29
 DB: 13 Gaps: 7

US-10-081-935-4 (1-194) x BU064419 (1-611)

QY 1 MetGlnPheSerHisAlaLeuValAlaGlyLeuAlaSerAlaGlnLeu 20
 DB 80 ATGAAGACTCCGTCGCTTCGTCGCTCCGCCGCTTCGCCGCTCACTCTT 139
 QY 21 ProAspIleProProCysAlaLeuAsnGlyPheValGlnAlaLeuGlyAsnAspGlyCys 40
 DB 140 GCCGACGTCCCAAGTCGCGCATCTTCCTCGACAGCCATCGCCAGCAGACGAGC 199
 QY 41 ThrArgLeuThrAspPheLysCysHisCysSerLys---ProGluLeuProGlyGlnIle 59
 DB 200 TGTGACAAAGCCACCTCGCTGTGTGTGCAAGGCTTCAGCGCTGCATCCATCCAGGCT 259
 QY 60 ThrProCysValGlnGluAlaCysProLeuAspAlaArgIleSerValSerAsnIleVal 79
 DB 260 ACCTCTTGTCATTGACGAGTGTGTAACGAC-----GTTGCCATCAACGAGGTC--- 310
 QY 80 ValAspGlnCysSerLysAlaGlyValProIleAspIleProProValAspThrThrAla 99
 DB 311 -----CTTCCCGCTACCGAAGACTCTGTC 334
 QY 100 AlaProGluProSerGluThr---AlaGluProThrAlaGluProThrGluGluProThr 118


```

Db      335 AAGAACCCCGCAAGAGTCTGAGGCCAAGTCCACCGCCGAG---GAGGAGAACCCACC 391
Qy      119 AlagluProthralagluProthr-----AlagluProthrhlsgluProthglu 135
          ||| ||| |||
Db      392 ACCACCCCGCTGCACCTTCGTCGTTGTCACACACACCGCCGAGGTTGTGAG 451
Qy      136 GluProthralavalProthrglyThrglyglYvalProthrglyThrglySerphe 155
          ||| ||| |||
Db      452 ACCACCCCGCTGCACACCTGCTGCTCCATTCATCCACACCGCTGCTGAGAG 511
Qy      156 ThrValThrglyAProthralaserThProAlagluPheProglYAlaglySerasn 175
          ||| ||| |||
Db      512 CCGCCACCAAGACACCCCGCGCTGCCACCCCTACCAAGAGGTCCGAGACGCCAAGCGT 571
Qy      176 ValArgAlaserValglYglYlIeAlAlAlaleuleuglyLeuAlAla 192
          ||| ||| |||
Db      572 -----GCTGCTGGCTCTCAAGGCTTCGGTGCC 598

RESULT 37
BU065692      611 bp  mRNA  linear  EST 26-AUG-2002
LOCUS         Fgr_8_E01_T3 Nitrogen-starved mycelia Gibberella zeae cDNA, mRNA
DEFINITION   sequence.
ACCESSION    BU065692
VERSION      BU065692.1 GI:22505981
KEYWORDS     EST.
SOURCE       Gibberella zeae
ORGANISM     Gibberella zeae
REFERENCE    1 (bases 1 to 611)
AUTHORS      Trail,F., Xu,J.-R., San Miguel,P., Halgen,R.G. and Kistler,H.C.
TITLE        Analysis of expressed sequence tags from Gibberella zeae (anamorph
JOURNAL      Fungal Genet. Biol. 38 (2), 187-197 (2003)
MEDLINE      22508120
PUBMED       12620255
COMMENT      Contact: Frances Trail
              Department of Plant Biology
              Michigan State University
              East Lansing, MI 48824, USA
              Tel: 517 432 2939
              Fax: 517 353 1926
              Email: trail@msu.edu
              Plate: 8 row: E column: 01.
              Location/Qualifiers
FEATURES
    source          1..611
                     /organism="Gibberella zeae"
                     /mol_type="mRNA"
                     /strain="NRRL 31084"
                     /db_xref="taxon:5518"
                     /clone_lib="Nitrogen-starved mycelia"
                     /note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      128 a 229 c 143 g 111 t
ORIGIN
Alignment Scores:
Pred. No.:      0.00188      Length:      611
Score:          164.50      Matches:      52
Percent Similarity: 40.61%      Conservative: 28
Best local Similarity: 26.40%      Mismatches: 88
Query Match:    16.08%      Indels:      29
DB:             13      Gaps:      7

US-10-081-935-4 (1-194) x BU065692 (1-611)
Qy      1 MetClnPheSerHsAlaleuIleAlaleuValAlaAlaglyLeuAlaserAlaglnu 20
          ||| ||| |||
Db      80 ATGAAGACTCGGCTTCTGCTTCCTCCCGCTGTTGCCGCTCAGGCTCAGCTCTTT 139
Qy      21 ProAspIleProProCysAlaleuAsnCysPheValGluAlaleuGlyAsnAspGlyCys 40
          ||| ||| |||
Db      140 GCCACGTCGCCAAGTGGCCATTCCTTCCTCGACAGGCCATTCGACGACGACGAC 199

```

```

Qy      41 ThrArgIleuThraPhePheCysHlsCysSerlys---ProgluleProglYglNile 59
          ||| ||| |||
Db      200 TGTGACAAGACCGAGCTTCGCTGTGTGTGCAAGGCTTCACCGCTGTTCATCCAAAGCT 259
Qy      60 ThrProCysValglGluAlaCysProleuAspAlaArgIleSerValIleVal 79
          ||| ||| |||
Db      260 ACCCTTGTGTGATTAAGAGAGTGTGTAACGAC-----GTTCGATCAACAGAGATC--- 310
Qy      80 ValAspGlnCysSerlysalaglyValproIleAspIleProProValAspThrThrala 99
          ||| ||| |||
Db      311 -----CTTCCGCTTACCGAACCTTCGC 334
Qy      100 AlaprogluProsergluThr-----AlagluProthralagluProthngluProthr 118
          ||| ||| |||
Db      335 AAGAACCCCGCAAGAGTCTGAGGCCAAGTCCACCGCCGAG---GAGGAGAACCCACC 391
Qy      119 AlagluProthralagluProthr-----AlagluProthrhlsgluProthglu 135
          ||| ||| |||
Db      392 ACCACCCCGCTGCACCTTCACCTTCGTCGTTGTCACACACCGCCGAGGTTGTGAG 451
Qy      136 GluProthralavalProthrglyThrglyglYvalProthrglyThrglySerphe 155
          ||| ||| |||
Db      452 ACCACCCCGCTGCACACCTGCTGCTCCATTCATCCACACCGCTGCTGAGAG 511
Qy      156 ThrValThrglyAProthralaserThProAlagluPheProglYAlaglySerasn 175
          ||| ||| |||
Db      512 CCGCCACCAAGACACCCCGCGCTGCCACCCCTACCAAGAGGTCCGAGACGCCAAGCGT 571
Qy      176 ValArgAlaserValglYglYlIeAlAlAlaleuleuglyLeuAlAla 192
          ||| ||| |||
Db      572 -----GCTGCTGGCTCTCAAGGCTTCGGTGCC 598

RESULT 38
AU284313      612 bp  mRNA  linear  EST 17-SEP-2002
LOCUS         AU284313 Dictyostelium discoidium FC (H.Urushihara) Dictyostelium
DEFINITION   discoidium cDNA clone FC-AK17 5', mRNA sequence.
ACCESSION    AU284313
VERSION      AU284313.1 GI:23034403
KEYWORDS     EST.
SOURCE       Dictyostelium discoidium
ORGANISM     Dictyostelium discoidium
REFERENCE    1 (bases 1 to 612)
AUTHORS      Muramoto,T., Suzuki,K., Shimizu,H., Kohara,Y., Koriki,E., Obara,S.,
              Tanaka,Y. and Urushihara,H.
TITLE        Generation and analysis of a gamete-specific gene pool in
              Dictyostelium discoidium (2002b)
JOURNAL      Unpublished
COMMENT      Contact: Hideko Urushihara
              Institute of Biological Sciences
              University of Tsukuba
              1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
              Tel: 81-298-53-4664
              Fax: 81-298-53-6614
              Email: hideko@biol.tsukuba.ac.jp.
              Location/Qualifiers
FEATURES
    source          1..612
                     /organism="Dictyostelium discoidium"
                     /mol_type="mRNA"
                     /strain="KAX3"
                     /db_xref="taxon:44689"
                     /clone_lib="FC-AK17"
                     /sex="mat A"
                     /cuesue_type="gamete"
                     /dev stage="Sexual"
                     /clone_lib="Dictyostelium discoidium FC (H.Urushihara)"
BASE COUNT      229 a 164 c 62 g 157 t
ORIGIN
Alignment Scores:
Pred. No.:      0.00206      Length:      612

```


Score:	164.00	Matches:	4
Percent Similarity:	45.26%	Conservative:	3
Best Local Similarity:	42.11%	Mismatches:	48
Query Match:	16.03%	Indels:	4
DB:	9	Gaps:	2

US-10-081-935-4 (1-194) X AU284313 (1-612)

[illegible]

RESULT 39	LOCUS	DEFINITION
BU061554	643 bp	mRNA
BU061554		linear
		EST 26-AUG-2002
		Fgf_10_D22_T3 Nitrogen-starved mycelia Gibberella zeae cDNA, mRNA

ACCESSION	GI:22501843
SEQUENCE	
BU061554	
BU061554.1	

SOURCE Gibberella zeae

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.

REFERENCE	JOURNAL
1 (pages 1 to 6-9)	
AUTHORS	
Titali, F., Xu, Y.-C., San Miguel, P., Halgren, R.G. and Kistler, H.C.	
TITLE	
Analysis of expressed sequence tags from <i>Gibberella zeae</i> (anamorph)	
<i>Fusarium graminearum</i>)	
Fungal Genet. Biol. 38 (2), 187-197 (2003)	

PUBMED 12620255
 COMMENT Contact: Frances Trail

Department of Plant Biology
Michigan State University
East Lansing, MI 48824, USA
Tel: 517 432 2939
Fax: 517 353 1926
Email: trall@msu.edu
Plate: 10 row: D column: 22.

FEATURES	Location/Qualifiers
source	1. .643

```

/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="NRRL 31084"
/db_xref="taxon:51518"
/clone_1fb="Nitrogen-starved mycelia"
/note="Vector: Uni-ZapIII; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      131 a      240 c      147 g      125 t
ORIGIN

```

Alignment Scores:	
Pred. No.:	0.00218
Score:	164.000
Percent Similarity:	42.428
Best Local Similarity:	26.778
Query Match:	16.038
DB:	13
Length:	643
Matches:	53
Conservative:	31
Mismatches:	90
Indels:	24
Gaps:	8

US-10-081-935-4 (1-194) X BU061554 (1-643)

QY 1 MetClnPheSerHisAlaLeuIleAlaLeuValAlaAlaGlyLeuAlaSerAlaGlnLeu 20

Db 90 ATGAGTACTCCGTTCGTTTCGTTCTCTCCCGCGTGTGCCCGCTACGGCTCAGTCTTT 145

```

Oy      21 ProAspIleProProCysAlaLeuAsnCysPheValGluAlaLeuGlyAsnAspGlyCys 40
      |||::||| |||||:: ||| ::|||:: ::::
Db      150 GCCACGCTCCCAAGTCCGCATCTTGCTTCGACAAAGGCATCGCCAGCGAACCAGC 209

```

41 ThrArgLeuThrAspPheLysCysHisCysSerLys--ProGluLeuProGlyGlnIle 59

Db 210 TGTGACAGACCGACTCCGCTGTCTGCAAGGCTTCAGCGCTGCCGATCCAGGCT 265

60 ThProCysValGluGluIleAlaCysProLeuSpAlaArgIleSerValSerAsnIleVal 79

Db 270 ACCTCTTGTCATTGACGAGTGTGTACCGAC-----GTTGCCATCAACGAGTC--- 320

80 ValAspGlnCysSerLysAlaGlyValProIleAspIleProProValAspThrThrAla 99

Db 321 -----CTTCCCGCTACCGAGAACCTCTGC 344

Qy 100 AlaProGluProSerGluThr---AlaGluProThrAlaGluProThrGluGluProThr 118

345 AAGAACCCCCCAAGAGCTGAGGCCAAGTCCACGCCGAG--GAGGAGAAGCCCCC 401

[illegible]

136 GluProThrAlaValProThrGluValValProThrGluValThrGluValSerPhe 155

Db 462 ACCACCGCGCTGCACCAACCACTGTCGTCCTCCATTCATTCACCAACCGCTGAGAG 521

156 ThrValThrGlyArgProThrAlaSerThrProAlaGluPheProGlyAlaGlySerAsn 175

Dp
522 CCCGCACAGCACCCCGCGGTCTACCCCTACCAAGGTTCCC---GAGCAGGCCAAC 578

QY 176 ValArgAlaSerValGlyGlyIleValAlaAlaLeuLeuGlyLeuAlaAlaTyr 193

Db 579 GGGCTGCTGGCCTCAAGGCTCGGTGCC-----CTGGCCATGGCTGCTTTC 626

RESULT 40
CD453301/c

LOCUS			
CD459301	874 bp	mRNA	linear
Fg08_03a11_A	Fg08_AAFC_ECORC	Fusarium graminearum	complex substrat
DEFINITION			
EST 03-JUN-200			

Accession
Gibberella zeae cDNA clone Fg08_03a11, mRNA sequence.
CD459301

KEYWORDS CD459301.1 GI:31374041
EST.
SOURCE C:\bbs\0115\2000

SOURCE
 ORGANISM
 Gibberella zeae
 Gibberella zeae
 Fungus: Fungi: Ascomycota: Dothideomycetes: Cordariaceae

[illegible]

AUTHORS
Watson, R. J., Heyes, R., Chapados, J., Couroux, P., Harris, L. J., Hattori, J., Lacroix, C., Queller, T., Robert, L. S., Singh, J. A., Sprott, D. and

Tinker, N.A.
A cDNA library prepared from *Fusarium graminearum* grown on a

complex plant substrate
unpublished
Journal
Comment

COMMENTARY
Contact: Watson, Robert J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada

Regulatory and Agricultural Canada
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA

Tel: (613) 759-1655
Fax: (613) 759-1701

FEATURES

Email: watsonj@agr.gc.ca
Location/Qualifiers

```
source
1. .874
/organism="Gibberella zeae"
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2003, 06:44:55 : Search time 1940 Seconds

(without alignments)
651.460 Million cell updates/sec

Title: US-10-081-935-1

Perfect score: 52
Sequence: 1 atcgagtcctcctcagctct.....cgtcgccgctcgcagctg 52

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vxl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	395	10	BF251279
2	52	100.0	401	10	BF252856
3	52	100.0	531	10	BF251334
4	52	100.0	557	10	BF252723

5	52	100.0	737	10	BF251103	BF251103	EST148360
6	39.2	75.4	311	13	BQ499065 <td>BQ499065<td>EST08290</td></td>	BQ499065 <td>EST08290</td>	EST08290
7	32.8	63.1	419	13	BQ761304 <td>BQ761304<td>EBR008_SO</td></td>	BQ761304 <td>EBR008_SO</td>	EBR008_SO
8	32.8	63.1	538	13	BQ767128 <td>BQ767128<td>EBR008_SO</td></td>	BQ767128 <td>EBR008_SO</td>	EBR008_SO
9	32.8	63.1	602	13	BQ767421 <td>BQ767421<td>EBR008_SO</td></td>	BQ767421 <td>EBR008_SO</td>	EBR008_SO
10	29.4	56.5	994	29	CNS06QTV <td>CNS06QTV<td>AL411113 T3 end of</td></td>	CNS06QTV <td>AL411113 T3 end of</td>	AL411113 T3 end of
11	27.4	52.7	559	29	B2235140 <td>B2235140<td>CH230-477</td></td>	B2235140 <td>CH230-477</td>	CH230-477
12	27.2	52.3	349	14	L38010 <td>L38010<td>BNMF0252 M</td></td>	L38010 <td>BNMF0252 M</td>	BNMF0252 M
13	26.8	51.5	1615	12	B1519263 <td>B1519263<td>603062017</td></td>	B1519263 <td>603062017</td>	603062017
14	26.6	51.2	559	14	CA766669 <td>CA766669<td>AF53-RP1</td></td>	CA766669 <td>AF53-RP1</td>	AF53-RP1
15	26.6	51.2	850	29	B2564779 <td>B2564779<td>pac82-164</td></td>	B2564779 <td>pac82-164</td>	pac82-164
16	26.4	50.8	1044	28	AQ743420 <td>AQ743420<td>HS 5387 B</td></td>	AQ743420 <td>HS 5387 B</td>	HS 5387 B
17	26.2	50.4	447	13	B066215 <td>B066215<td>HS01822*</td></td>	B066215 <td>HS01822*</td>	HS01822*
18	26.2	50.4	456	14	CB883923 <td>CB883923<td>HS07A16u</td></td>	CB883923 <td>HS07A16u</td>	HS07A16u
19	26.2	50.4	522	9	AL925722 <td>AL925722<td>AL925722</td></td>	AL925722 <td>AL925722</td>	AL925722
20	26.2	50.4	627	13	BQ466638 <td>BQ466638<td>HS01B22T</td></td>	BQ466638 <td>HS01B22T</td>	HS01B22T
21	26.2	50.4	1145	29	AG035039 <td>AG035039<td>Pan t10G1</td></td>	AG035039 <td>Pan t10G1</td>	Pan t10G1
22	26	50.0	885	14	CB680101 <td>CB680101<td>OSJNEF04G</td></td>	CB680101 <td>OSJNEF04G</td>	OSJNEF04G
23	26	50.0	1057	13	BU504141 <td>BU504141<td>AGENCOURT</td></td>	BU504141 <td>AGENCOURT</td>	AGENCOURT
24	25.8	49.6	353	28	BH225332 <td>BH225332</td> <td>1006107G1</td>	BH225332	1006107G1
25	25.8	49.6	803	29	BZ547247 <td>BZ547247<td>OGAKP227M</td></td>	BZ547247 <td>OGAKP227M</td>	OGAKP227M
26	25.8	49.6	873	29	BZ547241 <td>BZ547241<td>OGAKP227C</td></td>	BZ547241 <td>OGAKP227C</td>	OGAKP227C
27	25.6	49.2	123	12	B1189998 <td>B1189998<td>93d05f8.r</td></td>	B1189998 <td>93d05f8.r</td>	93d05f8.r
28	25.6	49.2	478	13	BU667306 <td>BU667306<td>U002E04 O</td></td>	BU667306 <td>U002E04 O</td>	U002E04 O
29	25.6	49.2	548	28	BH411461 <td>BH411461<td>1007022G0</td></td>	BH411461 <td>1007022G0</td>	1007022G0
30	25.6	49.2	630	10	AW984083 <td>AW984083<td>RC0-HN000</td></td>	AW984083 <td>RC0-HN000</td>	RC0-HN000
31	25.6	49.2	773	14	CB657824 <td>CB657824<td>OSJNEC13H</td></td>	CB657824 <td>OSJNEC13H</td>	OSJNEC13H
32	25.4	48.8	601	14	CA354617 <td>CA354617<td>626361 NC</td></td>	CA354617 <td>626361 NC</td>	626361 NC
33	25.2	48.5	402	14	CA642822 <td>CA642822<td>wre1n.pK0</td></td>	CA642822 <td>wre1n.pK0</td>	wre1n.pK0
34	25.2	48.5	522	14	CA712672 <td>CA712672<td>wdk3c.pK0</td></td>	CA712672 <td>wdk3c.pK0</td>	wdk3c.pK0
35	25	48.1	281	14	C72378 <td>C72378<td>C72378 C72378 Rice</td></td>	C72378 <td>C72378 C72378 Rice</td>	C72378 C72378 Rice
36	25	48.1	359	14	CA724779 <td>CA724779<td>wdf3f.pK0</td></td>	CA724779 <td>wdf3f.pK0</td>	wdf3f.pK0
37	25	48.1	449	9	AW730728 <td>AW730728<td>CA_Ba02</td></td>	AW730728 <td>CA_Ba02</td>	CA_Ba02
38	25	48.1	569	10	BE583376 <td>BE583376<td>11-9B-MY</td></td>	BE583376 <td>11-9B-MY</td>	11-9B-MY
39	25	48.1	591	12	BM370328 <td>BM370328<td>EBR008_SO</td></td>	BM370328 <td>EBR008_SO</td>	EBR008_SO
40	25	48.1	688	14	CB649291 <td>CB649291<td>OSJNEB13C</td></td>	CB649291 <td>OSJNEB13C</td>	OSJNEB13C
41	25	48.1	755	14	CB648531 <td>CB648531<td>OSJNEB11P</td></td>	CB648531 <td>OSJNEB11P</td>	OSJNEB11P
42	25	48.1	798	14	CB680316 <td>CB680316<td>OSJNEF04N</td></td>	CB680316 <td>OSJNEF04N</td>	OSJNEF04N
43	25	48.1	804	10	BF684156 <td>BF684156<td>963049F02</td></td>	BF684156 <td>963049F02</td>	963049F02
44	25	48.1	899	29	BZ574472 <td>BZ574472<td>msb2-3683</td></td>	BZ574472 <td>msb2-3683</td>	msb2-3683
45	25	48.1	1188	29	AG047770 <td>AG047770<td>Pan t10G1</td></td>	AG047770 <td>Pan t10G1</td>	Pan t10G1

ALIGNMENTS

RESULT 1
BF251279
LOCUS
DEFINITION
EST148360 Coccidioides immitis spherule cDNA library EST 15-NOV-2001
immitis cDNA clone C1A4F11 5' sequence, mRNA sequence.

ACCESSION
BF251279
VERSION
KEYWORDS
SOURCE
ORGANISM

Coccidioides immitis
Coccidioides immitis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE
1 (bases 1 to 395)
AUTHORS
Gardner, M.J. and Kirkland, T.
TITLE
Generation of ESTs from Coccidioides immitis spherule cDNA library
JOURNAL
Unpublished
COMMENT
Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@cigr.org

FEATURES
source
1..395
/organism="Coccidioides immitis"
/mol_type="mRNA"

/db_xref="taxon:5501"
/clone="CIAF11"
/dev_stage="spherule"
/lab_host="SOLR"
/clone_lib="Coccidioides immitis spherule cDNA library"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"
XhoI"
78 a 133 c 78 g 106 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 52; DB 10; Length 395;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCAGTCTCTCAGCCTCTCATCCCTCTGCTGCGCGGCTCGCCAGTG 52
|||||
142 ATGCAGTCTCTCAGCCTCTCATCCCTCTGCTGCGCGGCTCGCCAGTG 193
|||||

RESULT 2
BF252856 401 bp mRNA linear EST 15-NOV-2001
LOCUS
DEFINITION
immitis cDNA clone CIAB70 5' sequence, mRNA sequence.
ACCESSION
BF252856
VERSION
KEYWORDS
SOURCE
ORGANISM
Coccidioides immitis
Coccidioides immitis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; mitosporic Onygenales; Coccidioides.
1 (bases 1 to 401)
Gardner, M.J. and Kirkland, T.
Generation of ESTs from Coccidioides immitis spherule cDNA library
Unpublished
Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org.
Location/Qualifiers
1..401
/organism="Coccidioides immitis"
/mol_type="mRNA"
/db_xref="taxon:5501"
/clone="CIAB70"
/dev_stage="spherule"
/lab_host="SOLR"
/clone_lib="Coccidioides immitis spherule cDNA library"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"
77 a 149 c 89 g 86 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 52; DB 10; Length 401;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCAGTCTCTCAGCCTCTCATCCCTCTGCTGCGCGGCTCGCCAGTG 52
|||||
59 ATGCAGTCTCTCAGCCTCTCATCCCTCTGCTGCGCGGCTCGCCAGTG 110
|||||

RESULT 3
BF251334 531 bp mRNA linear EST 15-NOV-2001
LOCUS
DEFINITION
immitis cDNA clone CIAF68 5' sequence, mRNA sequence.
ACCESSION
BF251334
VERSION
KEYWORDS
EST.

SOURCE
ORGANISM
Coccidioides immitis
Coccidioides immitis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; mitosporic Onygenales; Coccidioides.
1 (bases 1 to 531)
Gardner, M.J. and Kirkland, T.
Generation of ESTs from Coccidioides immitis spherule cDNA library
Unpublished
Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org.
Location/Qualifiers
1..531
/organism="Coccidioides immitis"
/mol_type="mRNA"
/db_xref="taxon:5501"
/clone="CIAF68"
/dev_stage="spherule"
/lab_host="SOLR"
/clone_lib="Coccidioides immitis spherule cDNA library"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"
110 a 197 c 117 g 133 t

BASE COUNT

Query Match 100.0%; Score 52; DB 10; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCAGTCTCTCAGCCTCTCATCCCTCTGCTGCGCGGCTCGCCAGTG 52
|||||
137 ATGCAGTCTCTCAGCCTCTCATCCCTCTGCTGCGCGGCTCGCCAGTG 188
|||||

RESULT 4
BF252723 557 bp mRNA linear EST 15-NOV-2001
LOCUS
DEFINITION
immitis cDNA clone CIAA21 5' sequence, mRNA sequence.
ACCESSION
BF252723
VERSION
KEYWORDS
SOURCE
ORGANISM
Coccidioides immitis
Coccidioides immitis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; mitosporic Onygenales; Coccidioides.
1 (bases 1 to 557)
Gardner, M.J. and Kirkland, T.
Generation of ESTs from Coccidioides immitis spherule cDNA library
Unpublished
Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org.
Location/Qualifiers
1..557
/organism="Coccidioides immitis"
/mol_type="mRNA"
/db_xref="taxon:5501"
/clone="CIAA21"
/dev_stage="spherule"
/lab_host="SOLR"
/clone_lib="Coccidioides immitis spherule cDNA library"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"
110 a 197 c 117 g 133 t

BASE COUNT

ORIGIN

Query Match 100.0%; Score 52; DB 10; Length 557;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAGTTCTCTCAGCGCTCTCATCGCTCTCGCTGCGCGCCTCGCCAGTG 52
 |||||
 161 ATGCAGTTCTCTCAGCGCTCTCATCGCTCTCGCTGCGCGCCTCGCCAGTG 212

RESULT 5
 Bf251103 737 bp mRNA linear EST 15-NOV-2001
 LOCUS EST418360 Coccidioides immitis spherule cDNA library Coccidioides
 DEFINITION immitis cDNA clone CIAAC67 5' sequence, mRNA sequence.
 ACCESSION Bf251103
 VERSION Bf251103.1 GI:16931246
 KEYWORDS EST.
 SOURCE Coccidioides immitis
 ORGANISM Coccidioides immitis
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Onygenales; mitosporic Onygenales; Coccidioides.

REFERENCE 1 (bases 1 to 737)
 Gardner, M.J. and Kirkland, T.
 Generation of ESTs from Coccidioides immitis spherule cDNA library
 Unpublished
 Contact: Malcolm J. Gardner
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301 838 3519
 Fax: 301 838 0208
 Email: gardner@igr.org.

FEATURES
 source Location/Qualifiers
 1..737
 /organism="Coccidioides immitis"
 /mol_type="mRNA"
 /db_xref="taxon:5501"
 /clone="CIAAC67"
 /dev_stage="spherule"
 /lab_host="SOLR"
 /clone_1lb="Coccidioides immitis spherule cDNA library"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 136 a 271 c 168 g 162 t

ORIGIN

Query Match 100.0%; Score 52; DB 10; Length 737;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAGTTCTCTCAGCGCTCTCATCGCTCTCGCTGCGCGCCTCGCCAGTG 52
 |||||
 99 ATGCAGTTCTCTCAGCGCTCTCATCGCTCTCGCTGCGCGCCTCGCCAGTG 150

RESULT 6
 BQ499065 311 bp mRNA linear EST 31-OCT-2002
 LOCUS EST08290 Pb0001 Paracoccidioides brasiliensis cDNA, mRNA sequence.
 DEFINITION EST08290 Pb0001 Paracoccidioides brasiliensis cDNA, mRNA sequence.
 ACCESSION BQ499065
 VERSION BQ499065.1 GI:24452039
 KEYWORDS EST.
 SOURCE Paracoccidioides brasiliensis
 ORGANISM Paracoccidioides brasiliensis
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Onygenales; mitosporic Onygenales; Paracoccidioides.
 1 (bases 1 to 311)
 Goldman, G.H., Marques, E.R., Ribeiro, D.C.D., Bernardes, L.A., Puccia
 R., Travaassos, L.R., Nobrega, F.G., Nobrega, M.P., Savoldi-Barbosa, M.,
 Semighini, C.P. and Goldman, M.H.
 The Paracoccidioides brasiliensis EST genome project

TITLE

BASE COUNT 71 a 169 c 97 g 82 t

ORIGIN

JOURNAL
 CONTACT Eukaryotic Cell, (2002) In press
 Contact: Gustavo Henrique Goldman
 Laboratory of Molecular Biology
 Universidade de Sao Paulo - USP - FCBRP
 Av do Cafe S/N, CEP: 14040-903, Ribeirao Preto - SP, Brazil
 Email: ggoldman@usp.br.

FEATURES
 source Location/Qualifiers
 1..311
 /organism="Paracoccidioides brasiliensis"
 /mol_type="mRNA"
 /db_xref="taxon:121759"
 /clone_1lb="Pb0001"

BASE COUNT 73 a 88 c 59 g 91 t

ORIGIN

Query Match 75.4%; Score 39.2; DB 13; Length 311;
 Best Local Similarity 84.6%; Pred. No. 0.076;
 Matches 44; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGCAGTTCTCTCAGCGCTCTCATCGCTCTCGCTGCGCGCCTCGCCAGTG 52
 |||||
 91 ATGCAGTTCTCTCAGCGCTCTCATCGCTCTGTGTGCGCGCCTAGCAATG 142

RESULT 7
 B0761304 419 bp mRNA linear EST 26-JUL-2002
 LOCUS EBr006_S0001_A05_R root, 3 week, drought-stressed, cv Optic, EBr006
 DEFINITION Hordeum vulgare subsp. vulgare cDNA clone EBr006_S0001_A05 5', mRNA
 sequence.
 ACCESSION B0761304
 VERSION B0761304.1 GI:21969776
 KEYWORDS EST.
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 419)
 Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardie, L.,
 Ramsay, L., Machray, G., Marshall, D.F.M. and Maugh, R.
 Development of Barley Transcriptome Resources
 Unpublished
 Contact: Maugh R. Marshall DF
 Genome Dynamics/Computational Biology
 Scottish Crop Research Institute
 Invergowrie, Dundee, DD2 5DA, Scotland, UK
 Tel: 00 44 1382 562731
 Fax: 00 44 1382 562426
 Email: eel@scri.sari.ac.uk.

FEATURES
 source Location/Qualifiers
 1..419
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Optic"
 /db_xref="taxon:112509"
 /clone="EBr006_S0001_A05"
 /cissue_type="root"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_1lb="root, 3 week, drought-stressed, cv Optic,
 EBr006"
 /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
 Non-normalised library, directionally cloned into pSPORT1.
 Derived from roots of 3 week old drought stressed barley
 plants. After two weeks in control conditions the seedlings
 were raised above the hydroponic solution for 3 hours
 daily. N.B. This library has been superseded by EBr006.
 Developed as part of the barley transcriptome resources of
 BBR/C/SEERAD funded cereal IGF (Investigating Gene
 Function) project."

BASE COUNT 71 a 169 c 97 g 82 t

ORIGIN

Query Match 63.1%; Score 32.8; DB 13; Length 419;
 Best Local Similarity 76.9%; Pred. No. 5.8;
 Matches 40; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 1 ATGCAGTTCTCTCAGCCTCTCATCGCTCGCTGCGCGGCTCGCCAGTG 52
 27 ATGAAGTTCTCCACCACCTCTGATTCGCTTCGTTGCGCGGCTCGCCAGCG 78

Db 27 ATGAAGTTCTCCACCACCTCTGATTCGCTTCGTTGCGCGGCTCGCCAGCG 78

RESULT 8
 B0767128
 LOCUS
 DEFINITION B0767128 538 bp mRNA linear EST 26-JUL-2002
 EBr008_S0008_E05_R root, 3 week, drought-stressed, cv Optic, EBr008
 Hordium vulgare subsp. vulgare cDNA clone EBr008_S0008_E05_5', mRNA
 sequence.

ACCESSION
 VERSION B0767128 GI:21975602
 KEYWORDS
 SOURCE Hordium vulgare subsp. vulgare
 ORGANISM Hordium vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 538)
 Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardie, L.,
 Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
 Development of Barley Transcriptome Resources
 Unpublished
 Contact: Waugh R, Marshall DF
 Genome Dynamics/Computational Biology
 Scottish Crop Research Institute
 Invergowrie, Dundee, DD2 5DA, Scotland, UK
 Tel: 00 44 1382 562731
 Fax: 00 44 1382 562426
 Email: est@scri.sari.ac.uk.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1..538
 location/Qualifiers
 /organism="Hordium vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Optic"
 /db_xref="taxon:112509"
 /clone="EBr008_S0008_E05"
 /tissue_type="root"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_11b="root, 3 week, drought-stressed, cv Optic,
 EBr008"
 /note="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I;
 Non-normalised library, directionally cloned into pSPORT1.
 Derived from roots of 3 week old drought stressed barley
 plants. Developed as part of the barley transcriptome
 resources of BBSRC/SEERAD funded cereal IGF (Investigating
 Gene Function) project."

BASE COUNT 91 a 217 c 119 g 111 t

ORIGIN

Query Match 63.1%; Score 32.8; DB 13; Length 538;
 Best Local Similarity 76.9%; Pred. No. 6;
 Matches 40; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 1 ATGCAGTTCTCTCAGCCTCTCATCGCTCGCTGCGCGGCTCGCCAGTG 52
 56 ATGAAGTTCTCCACCACCTCTGATTCGCTTCGTTGCGCGGCTCGCCAGCG 107

Db 56 ATGAAGTTCTCCACCACCTCTGATTCGCTTCGTTGCGCGGCTCGCCAGCG 107

RESULT 9
 B0767421
 LOCUS
 DEFINITION B0767421 602 bp mRNA linear EST 26-JUL-2002
 EBr008_S0008_I22_R root, 3 week, drought-stressed, cv Optic, EBr008
 Hordium vulgare subsp. vulgare cDNA clone EBr008_S0008_I22_5', mRNA
 sequence.

ACCESSION
 VERSION B0767421
 KEYWORDS
 SOURCE Hordium vulgare subsp. vulgare
 ORGANISM Hordium vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 602)
 Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardie, L.,
 Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
 Development of Barley Transcriptome Resources
 Unpublished
 Contact: Waugh R, Marshall DF
 Genome Dynamics/Computational Biology
 Scottish Crop Research Institute
 Invergowrie, Dundee, DD2 5DA, Scotland, UK
 Tel: 00 44 1382 562731
 Fax: 00 44 1382 562426
 Email: est@scri.sari.ac.uk.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1..602
 location/Qualifiers
 /organism="Hordium vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Optic"
 /db_xref="taxon:112509"
 /clone="EBr008_S0008_I22"
 /tissue_type="root"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_11b="root, 3 week, drought-stressed, cv Optic,
 EBr008"
 /note="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I;
 Non-normalised library, directionally cloned into pSPORT1.
 Derived from roots of 3 week old drought stressed barley
 plants. Developed as part of the barley transcriptome
 resources of BBSRC/SEERAD funded cereal IGF (Investigating
 Gene Function) project."

BASE COUNT 100 a 251 c 129 g 122 t

ORIGIN

Query Match 63.1%; Score 32.8; DB 13; Length 602;
 Best Local Similarity 76.9%; Pred. No. 6.1;
 Matches 40; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 1 ATGCAGTTCTCTCAGCCTCTCATCGCTCGCTGCGCGGCTCGCCAGTG 52
 45 ATGAAGTTCTCCACCACCTCTGATTCGCTTCGTTGCGCGGCTCGCCAGCG 96

Db 45 ATGAAGTTCTCCACCACCTCTGATTCGCTTCGTTGCGCGGCTCGCCAGCG 96

RESULT 10
 B0767421
 LOCUS
 DEFINITION B0767421 994 bp DNA linear GSS 05-JUL-2001
 T3 end of clone AW0A009A12 of library AW0A from strain CLIB 89 of
 Yarrowia lipolytica, genomic survey sequence.

ACCESSION
 VERSION AL411113
 KEYWORDS
 SOURCE Yarrowia lipolytica
 ORGANISM Yarrowia lipolytica
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Dipodascaceae; Yarrowia.
 1 (bases 1 to 994)
 Soucier, J.L., Aigle, M., Artiguenave, F., Blandin, G.,
 Biotin-Fukuhara, M., Bon, E., Broctier, P., Casaregola, S.,
 de-Montigny, J., Dujon, B., Durren, P., Lepingle, A., Lorente, B.,
 Maupertuy, A., Neuvéglise, C., Ozier-Kalogeropoulos, O., Pottier, S.,
 Saurin, W., Tekala, F., Toffano-Nicolas, C., Wesolowski-Louvel, M.,
 Winkler, P. and Weisenbach, J.
 Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 yeast species for molecular evolution studies
 FEMS Lett. 487 (1), 3-12 (2000)

TITLE
 JOURNAL
 MEDLINE
 PUBMED

VERSION B0767421.1 GI:21975895
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1..602
 location/Qualifiers
 /organism="Hordium vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Optic"
 /db_xref="taxon:112509"
 /clone="EBr008_S0008_I22"
 /tissue_type="root"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_11b="root, 3 week, drought-stressed, cv Optic,
 EBr008"
 /note="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I;
 Non-normalised library, directionally cloned into pSPORT1.
 Derived from roots of 3 week old drought stressed barley
 plants. Developed as part of the barley transcriptome
 resources of BBSRC/SEERAD funded cereal IGF (Investigating
 Gene Function) project."

BASE COUNT 100 a 251 c 129 g 122 t

ORIGIN

Query Match 63.1%; Score 32.8; DB 13; Length 602;
 Best Local Similarity 76.9%; Pred. No. 6.1;
 Matches 40; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 1 ATGCAGTTCTCTCAGCCTCTCATCGCTCGCTGCGCGGCTCGCCAGTG 52
 45 ATGAAGTTCTCCACCACCTCTGATTCGCTTCGTTGCGCGGCTCGCCAGCG 96

Db 45 ATGAAGTTCTCCACCACCTCTGATTCGCTTCGTTGCGCGGCTCGCCAGCG 96

RESULT 10
 B0767421
 LOCUS
 DEFINITION B0767421 994 bp DNA linear GSS 05-JUL-2001
 T3 end of clone AW0A009A12 of library AW0A from strain CLIB 89 of
 Yarrowia lipolytica, genomic survey sequence.

ACCESSION
 VERSION AL411113
 KEYWORDS
 SOURCE Yarrowia lipolytica
 ORGANISM Yarrowia lipolytica
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Dipodascaceae; Yarrowia.
 1 (bases 1 to 994)
 Soucier, J.L., Aigle, M., Artiguenave, F., Blandin, G.,
 Biotin-Fukuhara, M., Bon, E., Broctier, P., Casaregola, S.,
 de-Montigny, J., Dujon, B., Durren, P., Lepingle, A., Lorente, B.,
 Maupertuy, A., Neuvéglise, C., Ozier-Kalogeropoulos, O., Pottier, S.,
 Saurin, W., Tekala, F., Toffano-Nicolas, C., Wesolowski-Louvel, M.,
 Winkler, P. and Weisenbach, J.
 Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 yeast species for molecular evolution studies
 FEMS Lett. 487 (1), 3-12 (2000)

TITLE
 JOURNAL
 MEDLINE
 PUBMED

```

REFERENCE
AUTHORS      2 (bases 1 to 994)
              Casaregola,S., Neuveglise,C., Lepingle,A., Bon,E., Feynerol,C.,
              Artiguenave,F., Wincker,P. and Gaillardin,C.
TITLE        Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia
              lipolytica
JOURNAL      FEBS Lett. 487 (1), 95-100 (2000)
MEDLINE      20584727
PUBMED       11152892
REFERENCE    3 (bases 1 to 994)
AUTHORS      Direct Submission
TITLE        Genoscope.
JOURNAL      Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
              2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
              seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT      This GSS is part of a random genomic sequencing program of thirteen
              yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
              exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
              Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
              lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
              angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
              Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
              5 kb were prepared and both extremities were sequenced. See
              keywords for description of this sequence and for the sequence of
              the other extremity of this insert.
FEATURES     Location/Qualifiers
              source
                1..994
                /organism="Yarrowia lipolytica"
                /mol_type="genomic DNA"
                /strain="CLIB 89"
                /db_xref="taxon:4952"
                /clone="AM0AA009A12"
                /clone_lib="AM0AA"
                /note="end : T3"
              misc_feature
                <653..>969
                /note="similar to Saccharomyces cerevisiae ORF YPL123C [
                similarity to ribonucleases ]
                1 putative frameshift(s)"
                /evidence=not_experimental
BASE COUNT   234 a 325 c 210 g 221 t 4 others
ORIGIN

Query Match      56.5%; Score 29.4; DB 29; Length 994;
Best Local Similarity 76.6%; Pred. No. 65;
Matches 36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 ATGCAGTTCTCTCAGCTCTCATCGCTCTCGTCGTCGCCGCCCTCGC 47
Db 471 ATGCAGTTCTCTCTTGCACCATCGCTACCGTCGCTGCCGCCGCTCTC 517

RESULT 11
BZ235140
LOCUS        BZ235140 559 bp DNA linear GSS 12-OCT-2002
DEFINITION   CH230-477F24.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
              CH230-477F24, genomic survey sequence.
ACCESSION    BZ235140
VERSION      BZ235140.1 GI:23893919
KEYWORDS     GSS.
SOURCE       Rattus norvegicus (Norway rat)
ORGANISM     Eukaryota; Chordata; Chordata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE    1 (bases 1 to 559)
AUTHORS      Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
              A., Gebregorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
              Jong,P. and Fraser,C.M.
JOURNAL      Rat BAC End Sequences from Library CHORI-230 MboI segment
              Unpublished
COMMENT      Other GSSs: CH230-477F24.TV
              Contact: Shaying Zhao
              Department of Eukaryotic Genomics
              The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
Plate: 477 row: F column: 24
Seq primer: SP6
Class: BAC ends.
FEATURES     Location/Qualifiers
              source
                1..559
                /organism="Rattus norvegicus"
                /mol_type="genomic DNA"
                /strain="BN/SSNHsd/MCW"
                /db_xref="taxon:10116"
                /clones="CH230-477F24"
                /sex="Female"
                /cell_type="Brain"
                /clone_lib="CHORI-230 Segment 2"
                /note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
                CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
                Pieter de Jong"
BASE COUNT   144 a 143 c 116 g 156 t
ORIGIN

Query Match      52.7%; Score 27.4; DB 29; Length 559;
Best Local Similarity 75.6%; Pred. No. 2.3e+02;
Matches 34; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 4 CAGCTCTCTCAGCTCTCATCGCTCTCGTCGTCGCCGCCCTCGCC 48
Db 25 CAGCTCTCTCAGCTCTCACTTCTCAAGTCTCTTACCTGCTCACC 69

RESULT 12
L38010
LOCUS        BNAF0225E Mustard flower buds Brassica rapa cDNA clone F0225, mRNA
DEFINITION   BNAF0225E Mustard flower buds Brassica rapa cDNA clone F0225, mRNA
              sequence.
ACCESSION    L38010
VERSION      L38010.1 GI:887186
KEYWORDS     EST.
SOURCE       Brassica rapa
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
              ; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE    1 (bases 1 to 349)
AUTHORS      Lim,C.O., Kim,H.Y., Kim,M.G., Lee,S.I., Chung,W.S., Park,S.H.,
              Hwang,I. and Cho,M.J.
TITLE        Expressed sequence tags of Chinese cabbage flower bud cDNA
JOURNAL      Plant Physiol. 111 (2), 577-588 (1996)
MEDLINE      96255495
PUBMED       8787028
COMMENT      Contact: Lim,C.O., Kim,M.G., Hwang,I. and Cho,M.J.
              Plant Molecular Biology and Biotechnology Research Center
              Email: pmbrc@nongae.gsnu.ac.kr.
FEATURES     Location/Qualifiers
              source
                1..349
                /organism="Brassica rapa"
                /mol_type="mRNA"
                /strain="pekinensis"
                /db_xref="taxon:3711"
                /clone="F0225"
                /clone_lib="Mustard flower buds"
                /note="Devel_stage = flower bud"
BASE COUNT   49 a 110 c 67 g 123 t
ORIGIN

```


University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

source
1. .850
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_432"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun library."
167 a 299 c 221 g 163 t

BASE COUNT

ORIGIN

Query Match 51.2%; Score 26.6; DB 29; Length 850;
Best Local Similarity 71.4%; Pred. No. 4.1e+02;
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 4 CAGTCTCTCAGCGCTCATCGCTCTCGTCTCGCGCGCTCGCGCAGTG 52
Db 434 CAGCTCGCTCAGCGCTCTCGTCTCGTCTCGCGCGCGCGCAGTG 386

RESULT 16

AQ743420/c
LOCUS
DEFINITION
HS_5387_B1_E08_SP6 RPCI-11 Human Male BAC Library Homo sapiens
Genomic Clone Plate=963 Col=15 Row=J, genomic survey sequence.

ACCESSION

Q743420.1 GI:5520942

KEYWORDS

SOURCE

ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1044)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D., and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
10449764

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 963 Row: J Column: 15
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 1044.

FEATURES

source
1. .1044
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate=963 Col=15 Row=J"
/sex="male"

/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"

BASE COUNT 295 a 160 c 411 g 148 t 30 others
ORIGIN
Query Match 50.8%; Score 26.4; DB 28; Length 1044;
Best Local Similarity 69.2%; Pred. No. 4.9e+02;
Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 1 ATGCAGTCTCTCAGCGCTCATCGCTCTCGTCTCGCGCGCTCGCGCAGTG 52
Db 725 ATTCATCTCTCAGACCCGCTGCTCTAGTCTCTTCTCTCCCTCACAGTG 674

RESULT 17

BQ662215/c
LOCUS
DEFINITION
HS01B22W HS Hordeum vulgare subsp. vulgare cDNA clone HS01B22
3-PRIME, mRNA sequence.

ACCESSION

Q662215 GI:21803348

KEYWORDS

SOURCE

ORGANISM
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.

REFERENCE

AUTHORS
Zhang,H., Potokina,E., Michalek,W., Weschke,W., Stein,N. and Graner
A.

TITLE

Barley ESTs from germinating seeds

JOURNAL

COMMENT

Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 447 Std Error: 0.00
Plate: 1 row: B column: 22

Seq Primer:

T7

Location/Qualifiers

1. .447
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/db_xref="GABI:142800"
/db_xref="taxon:112509"
/clone="HS01B22"
/tissue_type="embryo + scutellum"
/dev_stage="0-16 hours after imbibition"
/lab_host="XL10-Gold"
/clone_lib="HS"
/note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning
artefact caused by the kit, in most cases the EcoRI site
is NOT present, as well as the EcoRI adapter used for
cloning. To excise the insert, restriction sites upstream
EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
due to the cloning system used Blue/white selection for
recombinats is not 100% reliable."
93 a 137 c 164 g 53 t

BASE COUNT

ORIGIN

Query Match 50.4%; Score 26.2; DB 13; Length 447;
Best Local Similarity 72.3%; Pred. No. 4.9e+02;
Matches 34; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Best Local Similarity	75.6†	Pred. No. 5.7e+02;		
Matches	31; Conservative	0; Mismatches	10; Indels	0; Gaps

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1057)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM14044 row: b column: 15
High quality sequence stop: 397.

FEATURES

source

1..1057
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6491918"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."
BASE COUNT 156 a 379 c 267 g 253 t 2 others
ORIGIN

Query Match 50.0%; Score 26; DB 13; Length 1057;
Best Local Similarity 70.0%; Pred. No. 6.4e+02;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 2 TGCAGTTCTTCACGCTCTCATCGCTCTCGCTCGCGCGCGCTCGCCAGT 51
Db 693 TTCAGTTCCACGCTCTCTCTCGCTCTTTTCGCTCCCGCGCCAGT 732

RESULT 24
BH222532/c
LOCUS
DEFINITION

1006107G12.x1 1006 - RescueMu Grid G Zea mays genomic, genomic
survey sequence.

ACCESSION
BH222532
VERSION
BH222532.1 GI:16817477
KEYWORDS
SOURCE

ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 353)
Walbot,V.
Walbot,V.
Maize genomic sequences found using engineered RescueMu transposon
Unpublished
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 1006107 row: 12
Class: transposon-tagged.
Location/Qualifiers
1..353
/organism="Zea mays"
/mol_type="genomic DNA"
/cultiivar="mixed background W23/A188/B73"

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..353
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/cultiivar="mixed background W23/A188/B73"

Query Match 49.6%; Score 25.8; DB 29; Length 803;
Best Local Similarity 81.1%; Pred. No. 7e+02;
Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1006 - RescueMu Grid G"
/note="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site_1: BamHI; Site_2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
extracted from leaf punches, double digested using BamHI
and BglII, and ligated to form circular plasmids. DH10B
cells were transformed and then screened on LB plates with
ampicillin."
BASE COUNT 56 a 92 c 171 g 34 t
ORIGIN

Query Match 49.6%; Score 25.8; DB 28; Length 353;
Best Local Similarity 81.1%; Pred. No. 6.2e+02;
Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 9 CTCATCGCTCTCATCGCTCTCGCTCGCTCGCGCGCTC 45
Db 75 CGCTCGCGCTCTCTCGCTCTCGCTCGCGCGCTC 39

RESULT 25

BZ547247/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT


```

ORIGIN
Query Match      49.2%; Score 25.6; DB 13; Length 478;
Best Local Similarity 77.5%; Pred. No. 7.4e+02;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Oy  9 CTCTCAGCTCTCATCGCTCTCGTCGCTGCGCGCTCGCC 48
Db  226 CGCTCATGCTCTCATGCGCTCTCGTCGCGCGCGCGCC 265

RESULT 29
BH411461
LOCUS      1007022G03.x2 1007 - RescueMu Grid H Zea mays genomic, genomic
DEFINITION survey sequence.
ACCESSION  BH411461
VERSION     BH411461.1 GI:17582978
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 548)
AUTHORS   Walbot,V.
TITLE     Maize genomic sequences found using engineered RescueMu transposon
JOURNAL   Unpublished
COMMENT   Contact: Walbot V
          Department of Biological Sciences
          Stanford University
          855 California Ave, Palo Alto, CA 94304, USA
          Tel: 650 723 2227
          Fax: 650 725 8221
          Email: walbot@stanford.edu
          Plate: 1007022 column: 11
          Class: transposon-tagged.
FEATURES   Location/Qualifiers
            source          1..548
                        /organism="Zea mays"
                        /mol_type="genomic DNA"
                        /cultivar="mixed background W23/A188/B73"
                        /db_xref="taxon:4577"
                        /tissue_type="leaf"
                        /dev_stage="adult"
                        /lab_host="DH10B"
                        /clone_lib="1007 - RescueMu Grid H"
                        /note="Organ: leaf; Vector: RescueMu (engineered from
                        pBlueScript backbone); Site 1: BamHI; Site 2: BglII;
                        RescueMu is a 4.9 kb, modified maize Mu transposon
                        designed to allow plasmid rescue from total genomic DNA.
                        Mu elements insert preferentially into transcription
                        units. For more information on RescueMu, go to the web
                        site 'www.zmdb.iastate.edu' and follow the links for
                        'RescueMu.' Grid H was grown at Berkeley in 2001. DNA
                        was extracted from leaf punches, double digested using
                        BamHI and BglII, and ligated to form circular plasmids.
                        DH10B cells were transformed and then screened on LB
                        plates with ampicillin."
BASE COUNT  100 a 171 c 199 g 78 t
ORIGIN

Query Match      49.2%; Score 25.6; DB 28; Length 548;
Best Local Similarity 70.8%; Pred. No. 7.6e+02;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Oy  2 TGCAGTTCTCTCAGCTCTCATCGCTCTCGTCGCTGCGCGCTCGCCA 49
Db  364 TGCAGTCTCCACCGCCATCGCTCTCGTCGCTGCGCGCGCGCGCCA 411

RESULT 30
AW984083/c

```

```

LOCUS      AW984083      630 bp      mRNA      linear      EST 02-JUN-2000
DEFINITION  RC0-HN0006-160300-011-h09 HN0006 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW984083
VERSION     AW984083.1 GI:8175681
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 630)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
          Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
          Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
          Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
          ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
          Simpson,A.J.
TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
PUBMED     10737800
COMMENT    Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2=RC0-HN0006-160
          300-011-h09&t3=2000-03-16&t4=1)
          Seq primer: puc 18 forward
          High quality sequence start: 5
          High quality sequence stop: 478.
FEATURES   Location/Qualifiers
            source          1..630
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /dev_stage="Adult"
                        /clone_lib="HN0006"
                        /note="Organ: head normal; Vector: puc18; Site_1: SmaI;
                        Site_2: SmaI; A mini-library was made by cloning products
                        derived from ORESTES PCR (U.S. Letters Patent application
                        No. 196,716 - Ludwig Institute for Cancer Research)
                        profiles into the pUC 18 vector. Reverse transcription of
                        tissue mRNA and cDNA amplification were performed under
                        low stringency conditions."
BASE COUNT  148 a 192 c 187 g 103 t
ORIGIN

Query Match      49.2%; Score 25.6; DB 10; Length 630;
Best Local Similarity 70.8%; Pred. No. 7.7e+02;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Oy  4 CAGTTCTCTCAGCTCTCATCGCTCTCGTCGCTGCGCGCTCGCCAGT 51
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  234 CACTTTGCTCACCGCGTGGACGATCCGCTCCCTGCGCGCTCGCCCGT 187
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 31
CB657824
LOCUS      OSJNEC13H04.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA
DEFINITION  clone OSJNEC13H04 5', mRNA sequence.
ACCESSION  CB657824
VERSION     CB657824.1 GI:29661549
KEYWORDS   EST.
SOURCE     Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 773)
Jantassuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 13 row: H column: 04
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
1..773
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEC13H04"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEC"
/notes="vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"
BASE COUNT 139 a 265 c 240 g 129 t
ORIGIN

Query Match 49.2%; Score 25.6; DB 14; Length 773;
Best Local Similarity 77.5%; Pred. No. 8e+02;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 9 CTCCTACGCTCTCATCGCTCTCGCTCGCGCGCTCGCC 48
| ||||| ||||| ||||| ||||| ||||| |||||
Db 173 CGCTCATGCTCTCATCGCTCTCGCTCGCGCGCGCGCC 212

RESULT 32
CA354617 601 bp mRNA linear EST 05-NOV-2002
LOCUS
DEFINITION
626361 NCCCA 1RT Oncorhynchus mykiss cDNA clone 1RT78D19_B B10 5',
mRNA sequence.
ACCESSION
CA354617
VERSION
KEYWORDS
SOURCE
ORGANISM
Oncorhynchus mykiss (rainbow trout)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 601)
Rexroad,C.E. and Keele,J.W.
Sequence analysis of a rainbow trout normalized cDNA library
Unpublished
Contact: Rexroad CE
USDA, ARS, National Center for Cool and Cold Water Aquaculture
11876 Leetown Road, Kearneysville, WV 25430, USA
Tel: 304 724 8340 x2129
Fax: 304 725 0351
Email: crexroad@nccwa.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified by
cross match v0.990329.
Seq primer: AGCGGATACCAATTTCACACGGA.

```

```

FEATURES
source
Location/Qualifiers
1..601
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:9022"
/clone="1RT78D19_B B10"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="NCCCA 1RT"
/notes="vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from brain, gill, liver,
spleen, muscle, and kidney."
BASE COUNT 132 a 166 c 131 g 172 t
ORIGIN

Query Match 48.8%; Score 25.4; DB 14; Length 601;
Best Local Similarity 68.6%; Pred. No. 8.8e+02;
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 2 TGCAGTTCTCTCAGGCTCTCATCGCTCTCGCTCGCGCTCGCCAGTG 52
| ||||| ||||| ||||| ||||| ||||| |||||
Db 266 TGCCTTGTCTCACACTTACCTCTCTCTCCAGGGCGGCATGACTGTG 316
| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 33
CA642822 402 bp mRNA linear EST 23-NOV-2002
LOCUS
DEFINITION
wreln.pk0061.b12 wreln Triticum aestivum cDNA clone
wreln.pk0061.b12 5' end, mRNA sequence.
ACCESSION
CA642822
VERSION
KEYWORDS
SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
1 (bases 1 to 402)
Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
Miao,G., Caraher,N. and Hanafe,M.K.
DuPont Wheat cDNA Sequence
Unpublished
Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 13714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: M13.
Location/Qualifiers
1..402
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wreln.pk0061.b12"
/tissue_type="root"
/clone_lib="wreln"
/notes="vector: pBluescript SK+; Site 1: EcoRI; Site 2:
XhoI; Wheat (Triticum aestivum L.) root; normalized from
wreln library"
BASE COUNT 126 a 85 c 106 g 69 t 16 others
ORIGIN

Query Match 48.5%; Score 25.2; DB 14; Length 402;
Best Local Similarity 75.0%; Pred. No. 9.5e+02;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 7 TTCTCTCAGGCTCTCATCGCTCTCGCTCGCGCGGCTCG 46
| ||||| ||||| ||||| ||||| ||||| |||||
Db 210 TTCTTTNCTTCGACCATCGCTCTCTCTCTGCGGCTCG 171
| ||||| ||||| ||||| ||||| ||||| |||||

```

```
RESULT 34
CA712672      522 bp  mRNA  linear  EST 26-NOV-2002
LOCUS      wdk3c.pk006.l16 wdk3c Triticum aestivum cDNA clone wdk3c.pk006.l16
DEFINITION  5' end, mRNA sequence.
ACCESSION  CA712672
VERSION    CA712672.1 GI:25434465
KEYWORDS   EST.
SOURCE     Triticum aestivum (bread wheat)
ORGANISM   Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
            ; Triticeae; Triticum.
REFERENCE  1 (bases 1 to 522)
AUTHORS   Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
            Miao,G., Caraher,N. and Hanafey,M.K.
TITLE     DuPont Wheat cDNA Sequence
JOURNAL   Unpublished
COMMENT   Contact: Scott V. Tingey
            Crop Genetics
            E. I. DuPont de Nemours and Company
            1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
            Tel: 302-631-2602
            Fax: 302-631-2607
            Email: Scott.V.Tingey@USA.dupont.com
            Seq primer: M13.
            Location/Qualifiers
                1..522
                /organism="Triticum aestivum"
                /mol_type="mRNA"
                /db_xref="taxon:4565"
                /clone="wdk3c.pk006.l16"
                /tissue_type="kernel"
                /lab_host="DH10B"
                /clone_lib="wdk3c"
                /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
                XhoI; Wheat (Triticum aestivum L.) developing kernel, 14
                days after anthesis."
BASE COUNT  58 a 243 c 138 g 80 t 3 others
ORIGIN
Query Match 48.5%; Score 25.2; DB 14; Length 522;
Best Local Similarity 71.7%; Pred. No. 9.8e+02;
Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Oy 3 GCAGTTCTCTCAGCTCTCATCGCTCTCGTCTCGCGCGCTCGCC 48
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 276 GCAGTGGGCTCCCTCTCTCGCGCGCTCATCGCGCGCGCTCGCC 321

RESULT 35
C72378      281 bp  mRNA  linear  EST 04-APR-2002
LOCUS      C72378 Rice panicle at flowering stage Oryza sativa (japonica
DEFINITION  cultivar-group) cDNA clone E1517_2A, mRNA sequence.
ACCESSION  C72378
VERSION    C72378.1 GI:2427915
KEYWORDS   EST.
SOURCE     Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 281)
AUTHORS   Sasaki,T. and Yamamoto,K.
TITLE     Rice cDNA from panicle at flowering stage
JOURNAL   Unpublished
COMMENT   Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
```

```
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
PROJECT = 'RGP'.
Location/Qualifiers
    1..281
    /organism="Oryza sativa (japonica cultivar-group)"
    /mol_type="mRNA"
    /cultivar="Nipponbare"
    /db_xref="taxon:39947"
    /clone="E1517_2A"
    /dev_stage="flowering stage"
    /clone_lib="Rice panicle at flowering stage"
    /note="Organ: panicle; Rice cDNA from panicle at flowering
    stage"
BASE COUNT  45 a 98 c 96 g 39 t 3 others
ORIGIN
Query Match 48.1%; Score 25; DB 14; Length 281;
Best Local Similarity 69.4%; Pred. No. 1.e+03;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy 1 ATGCAGTTCTCTCAGCTCTCATCGCTCTCGTCTCGCGCTCGCC 49
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 ATGCAGCGGGCTCTCGCGCTCTCGCTCTCTCGCGCGCGCGCC 110

RESULT 36
CA724779      359 bp  mRNA  linear  EST 26-NOV-2002
LOCUS      wds3f.pk001.f15 wds3f Triticum aestivum cDNA clone wds3f.pk001.f15
DEFINITION  5' end, mRNA sequence.
ACCESSION  CA724779
VERSION    CA724779.1 GI:25446575
KEYWORDS   EST.
SOURCE     Triticum aestivum (bread wheat)
ORGANISM   Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
            ; Triticeae; Triticum.
REFERENCE  1 (bases 1 to 359)
AUTHORS   Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
            Miao,G., Caraher,N. and Hanafey,M.K.
TITLE     DuPont Wheat cDNA Sequence
JOURNAL   Unpublished
COMMENT   Contact: Scott V. Tingey
            Crop Genetics
            E. I. DuPont de Nemours and Company
            1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
            Tel: 302-631-2602
            Fax: 302-631-2607
            Email: Scott.V.Tingey@USA.dupont.com
            Seq primer: M13.
            Location/Qualifiers
                1..359
                /organism="Triticum aestivum"
                /db_xref="taxon:4565"
                /clone="wds3f.pk001.f15"
                /tissue_type="seedling"
                /lab_host="DH10B"
                /clone_lib="wds3f"
                /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
                XhoI; Wheat 10 day old seedling full length without
                biotin"
BASE COUNT  75 a 134 c 60 g 75 t 15 others
ORIGIN
Query Match 48.1%; Score 25; DB 14; Length 359;
Best Local Similarity 69.4%; Pred. No. 1.1e+03;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy 1 ATGCAGTTCTCTCAGCTCTCATCGCTCTCGTCTCGCGCTCGCC 49
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 80 ATGGGTTTCTCCACCATTGCGCGCTGTCTATCGCGCGCGGTGCC 128
```



```

RESULT 37
AW730728/c
LOCUS
DEFINITION
  GA_Ea027103 Gossypium arboreum 7-10 dpa fiber library Gossypium
  arboreum cDNA clone GA_Ea027103, mRNA sequence.
ACCESSION
AW730728
VERSION
AW730728.1 GI:7628277
KEYWORDS
EST.
SOURCE
  Gossypium arboreum
  Gossypium arboreum
  Gossypium arboreum
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
  ; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE
  1 (bases 1 to 449)
  Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
  D., Wood, T.C., Leslie, A., and Wilkins, T.A.
  An integrated analysis of the genetics, development, and evolution
  of the cotton fiber
  Unpublished
  Contact: Wing RA
  Clemson University Genomics Institute
  Clemson University
  100 Jordan Hall, Clemson, SC 29634, USA
  Tel: 864 656 7288
  Fax: 864 656 4293
  Email: rwing@clemson.edu
  High quality sequence stop: 449.
  Location/Qualifiers
    1..449
    /organism="Gossypium arboreum"
    /mol_type="mRNA"
    /strain="AKA"
    /cultivar="8400"
    /db_xref="taxon:29729"
    /clone="GA_Ea027103"
    /tissue_type="Fibers isolated from bolls harvested 7-10
    dpa"
    /lab_host="E. coli"
    /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
    /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT
  115 a 78 c 184 g
ORIGIN
Query Match 48.1%; Score 25; DB 9; Length 449;
Best Local Similarity 75.6%; Pred. No. 1.1e+03;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 TCGAGTTCTCAGCGTCTCATCGCTCGTCGTCGTCGCGGC 42
Db | ||||| ||||| ||||| ||||| ||||| ||||| |||||
262 TTTCGGTTTCGACGCTCTCATCGCTGTTCTCGCTCCCGGC 222

RESULT 38
BE583376
LOCUS
DEFINITION
  11-9B-MY Paojaemy Phytophthora sojae cDNA, mRNA sequence.
ACCESSION
BE583376
VERSION
BE583376.1 GI:9834325
KEYWORDS
EST.
SOURCE
  Phytophthora sojae
  Phytophthora sojae
  Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
  Phytophthora.
  1 (bases 1 to 569)
REFERENCE
  Qutob, D., Hraber, P.T., Sobral, B.W.S. and Gijsen, M.
  Comparative analysis of expressed sequences in Phytophthora sojae
  Plant Physiol. 123 (1), 243-254 (2000)
  MEDLINE
  20267956
  PUBMED
  10806241
  Contact: Gijsen M
  Agriculture and Agri-Food Canada

```

```

1391 Sandford Street, London, Ontario, Canada N5V 4T3
Tel: 519 457 1470
Fax: 519 457 3997
Email: gijzenm@em.agr.ca.
  Location/Qualifiers
    1..569
    /organism="Phytophthora sojae"
    /mol_type="mRNA"
    /strain="race 2, strain P6497"
    /db_xref="taxon:67593"
    /dev_stage="mycelium"
    /lab_host="E. coli strain XL0LR"
    /clone_lib="Psojaemy"
    /note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; This
    cDNA library was constructed from polyA+ enriched mRNA
    from mycelium grown in liquid medium for 3 weeks at 25 C
    in the dark in a liquid synthetic medium of 2.4 g sucrose
    , 0.27 g asparagine, 10 mg cholesterol, 10 mg ascorbate,
    and 2 mg thiamine, per liter, plus salts and minerals.
    Complementary DNA was synthesized from mRNA using an
    XhoI-poly(dT) linker-primer. EcoRI adapters were ligated
    to the blunt-ended cDNA fragments and the products were
    digested with XhoI for directional cloning into lambda ZAP
    Express vector. This lambda library was amplified once
    using E. coli host strain XL1 Blue MRF+. Inserts were then
    subcloned by mass excision using ExAssist helper phage for
    conversion into phagemid vector pBK-CMV in E. coli host
    strain XL0LR. Sequenced using T3 primer: 5' ATT AAC CCT
    CAC TAA AGG GA 3'."
BASE COUNT
  127 a 185 c 170 g 87 t
ORIGIN
Query Match 48.1%; Score 25; DB 10; Length 569;
Best Local Similarity 69.4%; Pred. No. 1.1e+03;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 ATCGAGTTCTCAGCGTCTCATCGCTCGTCGTCGCGGCCTCGCCA 49
Db | ||||| ||||| ||||| ||||| ||||| ||||| |||||
44 ATGAAGCTCTCTGCTGCTGCTCTCGCGCTCGCGCGCGCTTGCCA 92

RESULT 39
BM370328
LOCUS
DEFINITION
  BM370328.2 GI:21951439
  Hordeum vulgare subsp. vulgare
ACCESSION
BM370328
VERSION
BM370328.2
KEYWORDS
EST.
SOURCE
  Hordeum vulgare subsp. vulgare
  Hordeum vulgare subsp. vulgare
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Triticeae; Hordeum.
  1 (bases 1 to 591)
REFERENCE
  Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
  Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
  Development of Barley Transcriptome Resources
  Unpublished
  On Jan 10, 2002 this sequence version replaced gi:18113718.
  Contact: Waugh R, Marshall DF
  Genome Dynamics/Computational Biology
  Scottish Crop Research Institute
  Invergowrie, Dundee, DD2 5DA, Scotland, UK
  Tel: 00 44 1382 562731
  Fax: 00 44 1382 562426
  Email: est@scri.sari.ac.uk
  All sequence has a Phred quality score of 20 or over
  Seq primer: M13 reverse.
  Location/Qualifiers
    1..591
    /organism="Hordeum vulgare subsp. vulgare"

```

```

/mol_type="mRNA"
/cultivar="Optic"
/db_xref="taxon:112509"
/clone="EBro08_SQ003_N19"
/tissue_type="root"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="root, 3 week, drought-stressed, cv Optic,
EBro08"
/note="vector: pSPORT1; Site 1: Sal I; Site 2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from roots of 3 week old drought stressed barley
plants. Developed as part of the barley transcritome
resources of BBSRC/SEERAD funded cereal IGF (Investigating
Gene Function) project."
BASE COUNT      114 a      248 c      108 g      121 t
ORIGIN

      48.1%; Score 25; DB 12; Length 591;
Query Match      69.4%; Pred. No. 1.1e+03;
Best Local Similarity      0; Mismatches 15; Indels 0; Gaps 0;
Matches 34; Conservative

Qy      1 ATGCAGTTCTCTCAGGCTCTCATCGCTCTCGTCCGCGCTCGGCTCGGCA 49
|||||
Db      84 ATGCAGTTCCGCGCGCTCTTGGCCCTCGCGCTCGGCTCGGCTCGGCA 132
|||||

RESULT 40
CB649291
LOCUS      OSJNEB13C08.f OSJNEB Oryza sativa (japonica cultivar-group) cDNA
DEFINITION      clone OSJNEB13C08 5', mRNA sequence.
ACCESSION      CB649291
VERSION      CB649291.1 GI:29644284
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehartoideae; Oryzaceae; Oryza.
1 (bases 1 to 688)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 13 row: C column: 08
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
1..688
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEB13C08"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEB"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (Che 86061
)"

```

```

BASE COUNT      98 a      264 c      221 g      105 t
ORIGIN

      48.1%; Score 25; DB 14; Length 688;
Query Match      69.4%; Pred. No. 1.2e+03;
Best Local Similarity      0; Mismatches 15; Indels 0; Gaps 0;
Matches 34; Conservative

Qy      1 ATGCAGTTCTCTCAGGCTCTCATCGCTCTCGTCCGCGCTCGGCTCGGCA 49
|||||
Db      70 ATGCAGCGGGCTCTCGCGCTCTCGTCCGCTCTCGTCCGCGCGCGCA 118
|||||

```

Search completed: November 13, 2003, 08:51:58
Job time : 1949 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 13, 2003, 01:09:54 ; Search time 1945 Seconds
(without alignments)
1093.728 Million cell updates/sec

Title: US-10-081-935-1

Perfect score: 52

Sequence: 1 atgcagcttctcaagctctt.....cgctgcgcgcctgcagctg 52

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.ste.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.man.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	52	100.0	1201	8	CIU39835	U39835 Coccidioide
2	52	100.0	1234	8	CIU32518	U32518 Coccidioide
3	52	100.0	1435	8	U51200	U51200 Coccidioide
4	52	100.0	3821	8	AF013256	AF013256 Coccidioi
5	28.2	54.2	197059	2	AC105307	AC105307 Bos tauru
6	27.4	52.7	219080	2	AC120087	AC120087 Rattus no
7	27.4	52.7	253253	2	AC126527	AC126527 Rattus no
8	26.8	51.5	55939	2	AC091017	AC091017 Homo sapi
9	26.6	51.2	148301	8	AC098566	AC098566 Oryza sat
10	26.6	51.2	306067	8	AE017063	AE017063 Oryza sat
11	26.4	50.8	3582	8	AF092435	AF092435 Phaeospha
12	26.2	50.4	135357	2	AP004769	AP004769 Oryza sat
13	26.2	50.4	230357	2	AC140331	AC140331 Mus muscu
14	26	50.0	2595	3	LE1SER1HRE	LE1559 Leishmania
15	26	50.0	117846	2	AC073431	AC073431 Homo sapi
16	26	50.0	149082	2	AP005635	AP005635 Oryza sat
17	26	50.0	162208	2	AP005637	AP005637 Oryza sat
18	26	50.0	171665	2	AC084081	AC084081 Homo sapi
19	26	50.0	187336	2	AP005427	AP005427 Oryza sat
20	26	50.0	211291	9	AC107959	AC107959 Homo sapi
21	25.6	49.2	6000	4	OCU18102	Y18102 Oryctolagus
22	25.6	49.2	10951	1	AE008035	AE008035 Agrobacte
23	25.6	49.2	11764	1	AE009069	AE009069 Agrobacte
24	25.6	49.2	106732	8	AP005454	AP005454 Oryza sat
25	25.6	49.2	129487	2	AP003513	AP003513 Oryza sat
26	25.6	49.2	164978	2	AP003569	AP003569 Oryza sat
27	25.4	48.8	118444	2	AC114896	AC114896 Oryza sat
28	25.4	48.8	140556	2	AC130601	AC130601 Oryza sat
29	25.4	48.8	144077	2	AC099041	AC099041 Oryza sat
30	25.4	48.8	152359	2	CNS08C87	AL731758 Oryza sat
31	25.4	48.8	168864	2	CNS08CAF	AL831802 Oryza sat
32	25.4	48.8	292200	1	SC039129	AL939129 Streptomy
33	25.2	48.5	10037	1	AE005738	AE005738 Caulobact
34	25.2	48.5	141983	8	AP003047	AP003047 Oryza sat
35	25.2	48.5	281000	3	TBBCHRIA2	AL929604 Trypanoso
36	25	48.1	723	1	PSU96179	U96179 Pseudomonas
37	25	48.1	2898	1	AF232003	AF232003 Pseudomon
38	25	48.1	3076	1	AF461561	AF461561 Pseudomon
39	25	48.1	3719	1	MKFVCB	X98916 M.kandleri
40	25	48.1	10264	1	AE010324	AE010324 Methanopy
41	25	48.1	11867	1	AE005119	AE005119 Halobacte
42	25	48.1	14623	1	AE004898	AE004898 Pseudomon
43	25	48.1	101237	2	AC026810	AC026810 Homo sapi
44	25	48.1	114354	8	OSJN00227	AL731580 Oryza sat
45	25	48.1	117963	9	AC016651	AC016651 Homo sapi

ALIGNMENTS

RESULT 1
CIU39835
LOCUS
DEFINITION
Coccidioides immitis immunoreactive cell wall protein mRNA,
complete cds.
ACCESSION
U39835
VERSION
U39835.1
KEYWORDS
GI:1161373
ORGANISM
Coccidioides posadasii
Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; Mitosporic Onygenales; Coccidioidea;

CIU39835 1201 bp mRNA linear PLN 18-JAN-1996
Coccidioides immitis immunoreactive cell wall protein mRNA,
complete cds.

U39835
U39835.1
GI:1161373
Coccidioides posadasii
Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; Mitosporic Onygenales; Coccidioidea;
Duggar, K.O., Villareal, K.M., Ngyuen, A., Zimmermann, C.R., Law, J.H.
and Galgiani, J.N.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 253253)

REFERENCE AUTHORS

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, W., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Fallig, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, I., Garza, M., Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowitz, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Loulsegged, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokelimeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, K., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE

REFERENCE AUTHORS

2 (bases 1 to 253253)

REFERENCE AUTHORS

Worley, K. C.
Direct Submission
Submitted (06-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 253253)

REFERENCE AUTHORS

Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 13, 2003 this sequence version replaced gi:24819062. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GMEX

----- Summary Statistics
Center clone name: CH230-157M5

----- Assembly program: Atlas 3.0;

Consensus quality: 223138 bases at least Q40

Consensus quality: 225720 bases at least Q30

Consensus quality: 226900 bases at least Q20

Estimated insert size: 230245; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 250704: contig of 250704 bp in length

* 250705 250804: gap of unknown length

* 250805 251997: contig of 1193 bp in length

* 251998 252097: gap of unknown length

* 252098 253253: contig of 1156 bp in length.

FEATURES

source

1..253253
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-157M5"

misc_feature

1..1369
/note="wgs end_extension"
clone_end:T7"
3331..4610

misc_feature

/note="wgs end_extension"
clone_end:T7"
9685..10077

misc_feature

/note="clone_boundary"
clone_end:T7"
site:ECORI

misc_feature

end_sequence:BH286818"
complement(243817..244333)
/note="clone boundary"
clone_end:Sp6

misc_feature

site:ECORI
end_sequence:BH286821"
246750..248811
/note="wgs end_extension"
clone_end:Sp6"

misc_feature

248862..250704
/note="wgs end_extension"
clone_end:Sp6"

BASE COUNT 64432 a 49523 c 48940 g 64796 t 25562 others
ORIGIN

Query Match 52.7%; Score 27.4; DB 2; Length 253253;
Best Local Similarity 75.6%; Pred. No. 3e+02;
Matches 34; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

ORGANISM	Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartodeae; Oryzeae; Oryza. 1 (bases 1 to 306067)			mRNA		
REFERENCE AUTHORS CONSTRM TITLE	The Rice Chromosome 10 Sequencing Consortium In-depth view of structure, activity, and evolution of rice chromosome 10 Science 300, 1566-1569 (2003) 2 (bases 1 to 306067) Buell,C.R., Wing,R.A., McCombie,W.R., Messing,J. and Yuan,Q. Direct Submission Submitted (05-MAY-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA This is the pseudomolecule for rice chromosome 10, which was constructed by resolving discrepancies between overlapping BACs, trimming the overlap regions, and linking the unique sequences to form a contiguous sequence. Genes in individual BAC clone were identified by a combination of several methods: Gene prediction programs, searches of the complete sequence against a peptide database and EST databases. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).			CDS		
FEATURES source	1. 306067 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="genomic DNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="10" misc_feature <1. .100654 /note="Chromosome Sequence Derivation: nucleotide sequence in this region was derived from BAC clone OSJNBa0061H20 (GB:AC113337)." 3288. .>3647 /locus_tag="OSJNBa0061H20.7" /note="hypothetical protein" 3288. .>3647 /locus_tag="OSJNBa0061H20.7" 3288. .>3647 /locus_tag="OSJNBa0061H20.7" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52356.1" /db_xref="GI:31430447" /translations="MGFYPAATTTLSTAACVLVVGGLILARATVROEHGQRLPLPS GSGRRQLRLSSNYPAGNQLRRRLRRDRARHYWGHGDSAHLLGGGAIAIPDVLVGLDERCIKPAQCPSAS" 8040. .>11514 /locus_tag="OSJNBa0061H20.8" /note="hypothetical protein" join(<8040, .8132,8749, .8836,9821, .9901,10626, .10672, 11374, .>11514) /locus_tag="OSJNBa0061H20.8" join(8040, .8132,8749, .8836,9821, .9901,10626, .10672, 11374, .11514) /locus_tag="OSJNBa0061H20.8" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52357.1" /db_xref="GI:31430448" /translation="MAPAHGRRLASKGGSGNHALPSARSGGRAGAMGAKMGCTWLFLF QSDAGDMETCGSRGGLGLLRQAQSGCLRRMHPLPGASVMRGMQCPKNVISVGL FENYHRISINDVLWWSKLDSPPSVNTAISLLQGSDEKQAVIM" complement (<21422, .>28083) /locus_tag="OSJNBa0061H20.9" /note="Similar to RETROTRANSPOSABLE ELEMENT TP2 155 KDA			gene		
misc_feature				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJNBa0061H20.11" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52360.1" /db_xref="GI:31430451"		
mRNA				/translation="MDESSDLWRQGMALVGLPPAPVAVSAIVVFYLYLTWQMDYE EQLRRRTQAGLWLVLLVALVLLGSHALVDAGRVAVPVSWRWGGSGYSSGDDGG "		
CDS				complement (<34426, .>35324) /locus_tag="OSJNBa0061H20.10" /note="hypothetical protein" complement (join(<34426, .34543,34635, .34835,35161, .>35324)) /locus_tag="OSJNBa0061H20.10" complement (join(34426, .34543,34635, .34835,35161, .35324)) /locus_tag="OSJNBa0061H20.10" /codon_start=1 /product="hypothetical protein" /protein_id="AAP52359.1" /db_xref="GI:31430450"		
gene				/translation="MEAVARHLGHVLSAAPTNEHDCSVTLGPNMSGATSEARSLYVP PPLEPETLVALKLSQKASTFPDTHAAGRHPLHWDSPATQLELDNPLEEYGLIQ AAVGSNFSNSKEAVRGLATYTLSSRRVPASGDAEDFVPCDGHLMVGKEQEERG " complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" /note="hypothetical protein" complement (<36243, .>36653) /locus_tag="OSJNBa0061H20.11" complement (36243, .36653) /locus_tag="OSJ		

are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

- * NOTE: This is a 'working draft' sequence.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

FEATURES

source
1. .135357
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultiar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="P0017C12"
37722 a 28659 c 29003 g 39973 t

BASE COUNT

Query Match 50.4%; Score 26.2; DB 2; Length 135357;
Best Local Similarity 72.3%; Pred. No. 7e+02;
Matches 34; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 TGCAGTCTCTCAGGCTCATGCTCTCGTGGCTGGCGCTCGCC 48
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 35193 TGCAGTCTCTCAGGCTCATGCTCTCGTGGCTGGCGCTCGCC 35239
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13

AC140331

LOCUS AC140331 230357 bp DNA linear HTG 23-FEB-2003
DEFINITION Mus musculus chromosome UNK clone RP23-406J20, WORKING DRAFT

SEQUENCE, 7 unordered pieces.

ACCESSION

AC140331.1 GI:28475582

HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.

KEYWORDS

Mus musculus

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 230357)

McPherson,J.D. and Waterston,R.H.

The sequence of Mus musculus clone

Unpublished

2 (bases 1 to 230357)

McPherson,J.D. and Waterston,R.H.

Direct Submission

Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc/index.shtml

Contact: submissions@wustl.edu

----- Project Information -----

Center project name: M_BA0406J20

----- Summary Statistics -----

Sequencing vector: M13; 0%

Chemistry: Dye-terminator; 100%

Chemistry: Dye-terminator; 100%

Chemistry: Dye-terminator; 100%

Assembly program: Phrap; version 0.990319

Consensus quality: 229568 bases at least Q40

Consensus quality: 230223 bases at least Q30

Consensus quality: 230705 bases at least Q20

Insert size: 196000; agarose-fp

Insert size: 244718; sum-of-contigs

Quality coverage: 17.67 in Q20 bases; agarose-fp

Quality coverage: 11.05 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 7 contigs. The true order of the pieces

- * is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
- * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

FEATURES

source
1. .1255: contig of 1255 bp in length
1256: gap of unknown length
1356: contig of 20597 bp in length
21952: gap of unknown length
22052: gap of unknown length
52473: contig of 30421 bp in length
52573: gap of unknown length
52574: contig of 42538 bp in length
95111: gap of unknown length
95112: contig of 39840 bp in length
135051: contig of 35759 bp in length
135052: gap of unknown length
135151: contig of 59347 bp in length.
170910: gap of unknown length
170911: contig of 59347 bp in length.
171011: contig of 59347 bp in length.

Location/Qualifiers

1. .230357
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-406J20"

misc_feature

1. .1255

/note="assembly_name:Contig28"

misc_feature

1356. .21952

/note="assembly_name:Contig32"

misc_feature

22053. .52473

/note="assembly_name:Contig33"

misc_feature

52574. .95111

/note="assembly_name:Contig34"

misc_feature

95212. .135051

/note="assembly_name:Contig35"

misc_feature

135152. .170910

/note="assembly_name:Contig36"

misc_feature

171011. .230357

/note="assembly_name:Contig37"

BASE COUNT

63215 a 50684 c 50723 g 65054 t

ORIGIN

681 others

Query Match 50.4%; Score 26.2; DB 2; Length 230357;

Best Local Similarity 79.5%; Pred. No. 6.7e+02;

Matches 31; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGAGTCTCTCAGGCTCATGCTCTCGTGGCTGGCGCTCGCC 39

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 107371 ATGAGTCTCTCAGGCTCATGCTCTCGTGGCTGGCGCTCGCC 107409

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14

LEISERTHRE

LOCUS Leishmania chagasi (clone Cl) protein phosphatase-2C gene, complete

DEFINITION cds.

ACCESSION L15559

VERSION L15559.1 GI:293067

KEYWORDS protein phosphatase-2C; serine/threonine phosphatase.

SOURCE Leishmania donovani

ORGANISM Leishmania donovani

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Leishmania.

1 (bases 1 to 2595)

Burns,J.M. Jr., Parsons,M., Rosman,D.E. and Reed,S.G.

Molecular cloning and characterization of a 42-kDa protein

phosphatase of Leishmania chagasi

J. Biol. Chem. 268 (23), 17155-17161 (1993)

JOURNAL 93352496

MEDLINE 8394331

PUBMED

COMMENT Original source text: Leishmania donovani (individual isolate

MHOM/BR/82/BA-2, sub_species chagasi) promastigote DNA.

```
FEATURES
  source
1. .2595
  /organism="Leishmania donovani"
  /isolate="MHOW/BR/82/BA-2"
  /sub_species="chagasi"
  /db_xref="taxon:5661"
  /dev_stage="promastigote"
  /germline
891. .2111
  /function="serine/threonine phosphatase"
  /standard_name="PP2C"
  /codon_start=1
  /evidence=experimental
  /product="protein phosphatase-2C"
  /protein_id="AAA02864.1"
  /db_xref="GI:293068"
  /translation="MGIPLPKPVMTQLQERYGNAIFRCGSGNVGYRETMEADHLTYL
TDSMGFFGVDFGHVDCOSYLERAMSAIEKESIPMTDERMKELALRIDQEWMSGR
EGSGTGFFVALKEGNKVHLQVGNVDSRVVACIDGVCPLTEDHKPNNEGERORIEN
CAGRVENNRVDGSLAVRAGDREYKLGSSQLEQKVIADLVQHKFTFDSNDFVLL
CCDGVFGNFPNEEVAVYVKKQLETCNDLAEVAGRVCEEATERGSRDNI SCMIVQFKD
GSOYAAEPHTTVPGPFSAPRNSGFRKAYESMADKNTTGVALLERRYDTLKAAEALT
PEETEELSQFENGPEAKLTGAEROKWFSNVFKLCEAASNGPSQDMERLQSLQQQAGI
PLSTLLSLMGEQTQ"
BASE COUNT      555 a   797 c   687 g   556 t
ORIGIN
Query Match      50.0%; Score 26; DB 3; Length 2595;
Best Local Similarity 76.2%; Pred. No. 1.1e+03;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy  8  TCTCTACGCTCTATCGCTCTCTCGCTCGCGCGCTCGCCA 49
Db  493 TCGCTCTCGCTCTATCGATCCAGTCGCGCCCTCCCTCCCA 534

RESULT 15
AC073431
LOCUS      AC073431      117846 bp      DNA      linear      HTG 12-JUL-2001
DEFINITION Homo sapiens chromosome 8 clone GSI-146A9, WORKING DRAFT SEQUENCE,
6 unordered pieces.
AC073431
VERSION    AC073431.10 GI:14702002
KEYWORDS   HTG: HTGS PHASE1; HTGS_DRAFT.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 117846)
AUTHORS    McCombie,W.R.
TITLES     Human Genomic Sequence, Chromosome 8
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 117846)
AUTHORS    McCombie,W.R.
TITLES     Direct Submission
JOURNAL    Submitted (17-JUN-2000) Lita Annenberg Hazen Genome Sequencing
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
Harbor, NY 11724, USA
COMMENT    On Jul 12, 2001 this sequence version replaced gi:14423582.
----- Genome Center
Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
Laboratory
Center code: CSHL
Web site: http://www.cshl.org/genseq
Contact: mcombie@cshl.org
----- Project Information
Center project name: GSI-146A9
Center clone name: GSI-146A9
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 81715: contig of 81715 bp in length
* 81716 82745: gap of unknown length
* 82746 98894: contig of 16149 bp in length
* 98895 99924: gap of unknown length
* 106720 107749: contig of 6796 bp in length
* 107749 112805: contig of 5057 bp in length
* 112806 113834: gap of unknown length
* 113835 115554: contig of 1720 bp in length
* 115555 116583: gap of unknown length
* 116584 117846: contig of 1263 bp in length.
Location/Qualifiers
1. .117846
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="8"
  /clones="GSI-146A9"
  /clone_lib="Genome Systems 1"
BASE COUNT      31403 a   25767 c   26372 g   28909 t   5395 others
ORIGIN

Query Match      50.0%; Score 26; DB 2; Length 117846;
Best Local Similarity 70.0%; Pred. No. 8.1e+02;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy  1  ATGCAGTTCTCTACGCTCTCATCGCTCTCGTCGCTCGCGCTCGCCAG 50
Db  1416 ATGGAGTTCTCTGAACCTCTGCTGCTCTGAGGCTGCTGACTCTCGAG 1465

RESULT 16
AP005635
LOCUS      AP005635      149082 bp      DNA      linear      HTG 08-AUG-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 9 clone P0646C08,
*** SEQUENCING IN PROGRESS ***.
AC005635
VERSION    AP005635.1 GI:22138859
KEYWORDS   HTG: HTGS PHASE2.
SOURCE     Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE  1
AUTHORS    Sasaki,T., Matsumoto,T. and Katayose,Y.
TITLES     Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC
clone:P0646C08
JOURNAL    Published Only in Database (2002)
REFERENCE  2 (bases 1 to 149082)
AUTHORS    Sasaki,T., Matsumoto,T. and Katayose,Y.
TITLES     Direct Submission
JOURNAL    Submitted (07-AUG-2002) Takuji Sasaki, National Institute of
Agricultural Sciences, Rice Genome Research Program, Kamondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT    NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1. .149082
FEATURES
  source
```

```

/organism="Oryza sativa (japonica cultivar-group)"
/mol type="genomic DNA"
/cultivar="Nipponbare"
/db xref="taxon:39947"
/chromosome="9"
/clone="P0646C08"

BASE COUNT 42056 a 31796 c 31369 g 43559 t 302 others
ORIGIN
Query Match 50.0%; Score 26; DB 2; Length 149082;
Best Local Similarity 76.2%; Pred. No. 7.9e+02;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 9 CTCACGCTCTCATCGCTCTCGCTGCGCGCCCTGCGCCAG 50
   ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8117 CGCTCGGCTCTCTCGCTGCTTCCCGCGCTCTGCGCCG 8158

RESULT 17
AP005637 162208 bp DNA linear HTG 08-AUG-2002
LOCUS Oryza sativa (japonica cultivar-group) chromosome 9 clone P0711A01,
DEFINITION *** SEQUENCING IN PROGRESS ***.
ACCESSION AP005637 GI:22138861
VERSION HTG; HTGS_PHASE2.
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Enarthroideae; Oryzaceae; Oryza.
1
Sasaki, T., Matsumoto, T. and Katayose, Y.
Oryza sativa nipponbare (CA3) genomic DNA, chromosome 9, PAC
clone: P0711A01
Published Only in Database (2002)
2 (bases 1 to 162208)
Sasaki, T., Matsumoto, T. and Katayose, Y.
Direct Submission
Submitted (07-AUG-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kamondai
2-1-2 Teikuba Ibaraki 305-8602, Japan
E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
Location/Qualifiers
1. .162208
/organism="Oryza sativa (japonica cultivar-group)"
/mol type="genomic DNA"
/cultivar="Nipponbare"
/db xref="taxon:39947"
/chromosome="9"
/clone="P0711A01"

BASE COUNT 46642 a 33101 c 33667 g 48396 t 402 others
ORIGIN
Query Match 50.0%; Score 26; DB 2; Length 162208;
Best Local Similarity 76.2%; Pred. No. 7.9e+02;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 9 CTCACGCTCTCATCGCTCTCGCTGCGCGCCCTGCGCCAG 50
   ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 139467 CGCTCGGCTCTCTCGCTGCTTCCCGCGCTCTGCGCCG 139508

```

RESULT 18

AC084081/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC084081 171665 bp DNA linear HTG 27-MAY-2001
Homo sapiens chromosome 8 clone RP11-89M8 map 8, WORKING DRAFT
SEQUENCE, 9 unordered pieces.

AC084081.6 GI:14210547
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

Homo sapiens (human)
Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 171665)

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 8, clone RP11-89M8
Unpublished

2 (bases 1 to 171665)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bida, F., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,
DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Ferreira, P.,
FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRocque, K.,
Lamazares, R., Landers, T., Lehotzky, J., Levine, R., Lieu, C., Liu, G.,
Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,
McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
Morrow, J., Murphy, T., Navlyar, J., Norman, C. H., O'Connor, T.,
O'Donnell, P., O'Neil, D., Oliver, T. M., Oliver, J., Peterson, K.,
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,
Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,
Zimmer, A. and Zody, M.

Direct Submission

Submitted (12-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 27, 2001 this sequence version replaced gi:14196390.

All repeats were identified using RepeatMasker:

Smit, A. F. A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Center project name: L11307

Center clone name: 89_M_8

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 168994 bases at least Q40

Consensus quality: 169921 bases at least Q30

Consensus quality: 170318 bases at least Q20

Insert size: 169000; agarose-gel

Quality coverage: 9.9 in Q20 bases; agarose-gel

Quality coverage: 9.8 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently

consists of 9 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.


```

REFERENCE
1
Sasaki,T., Matsumoto,T. and Katayose,Y.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, PAC
clone:P0684H11
JOURNAL
Published Only in Database (2002)
REFERENCE
2 (bases 1 to 187336)
AUTHORS
Sasaki,T., Matsumoto,T. and Katayose,Y.
DIRECT SUBMISSION
Submitted (12-JUN-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasakienias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/.
Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
location/Qualifiers
1..187336
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="9"
/clone="P0684H11"
BASE COUNT 53235 a 38848 c 39766 g 55120 t 367 others
ORIGIN
Query Match 50.0%; Score 26; DB 2; Length 187336;
Best Local Similarity 76.2%; Pred. No. 7.8e+02;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 9 CTCACGCTCTCATCGCTCTCGTGGTCGCCGCCTCCAG 50
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 83156 CGCTCGGCTCTCTCGTGGTCGCCGCCTCCAG 83197
||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 20
LOCUS AC107959 211291 bp DNA linear PRI 28-OCT-2002
DEFINITION Homo sapiens chromosome 8, clone RP11-875011, complete sequence.
ACCESSION AC107959
VERSION AC107959.8 GI:24416019
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 211291)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 8, clone RP11-875011
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 211291)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Batra,N., Bastien,V., Boguslavsky,L., Bouckghalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczyk,J., Levine.R., Liu,G., MacLean.C.,
Macdonald.P., Major.J., Marquis.N., Matthews.C., McCarthy.M.,
McEwan.P., McKernan.K., Meldrim,J., Menesh.L., Minova.T.,
Menga.V., Murphy,T., Naylor,J., Nguyen.C., Nicol.R., Norbu,C.,
Norman.C.H., O'Connor,T., O'Donnell,P., O'Neill.D., Oliver,J.,

```



```

repeat_region      /rpt family="LTR36"
16283..16415
/rpt family="L2"
repeat_region      complement(16758..17011)
/rpt family="LTR37A"
17012..17044
/rpt family="T-rich"
repeat_region      complement(17045..17191)
/rpt family="LTR37A"
17545..17838
/rpt family="AluSx"
repeat_region      complement(18260..18562)
/rpt family="AluJo"
repeat_region      complement(18764..18965)
/rpt family="L1MD3"
18966..19348
/rpt family="MSTA"
repeat_region      complement(19349..19654)
/rpt family="AluSp"
19655..19695
/rpt family="MSTA"
repeat_region      complement(19696..19835)
/rpt family="L1MD3"
repeat_region      complement(19829..20084)
/rpt family="L1MC4a"
repeat_region      complement(20255..20878)

Query Match      50.0%; Score 26; DB 9; Length 211291;
Best Local Similarity 70.0%; Pred. No. 7.7e+02;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 ATGAGTCTCTCAGCTCTCATCGCTCTCTGCTCGCTCGGCTCGGCTCGAG 50
|||||
Db 24135 ATGAGTCTCTGAACCTCTGCTGCTCTGAGGCTCGCTGACTCTCGAG 24184
|||||

RESULT 21
OCU18102/c
LOCUS      Oryctolagus cuniculus mRNA for titin.
DEFINITION Y18102
ACCESSION Y18102.1 GI:3928488
VERSION titin; titin gene.
KEYWORDS Oryctolagus cuniculus (rabbit)
SOURCE Oryctolagus cuniculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1
AUTHORS Gregorio,C.C., Trombitas,T., Kolmerer,B., Stier,G., Granzier,H.,
Kunke,K., Suzuki,K., Obermayr,F., Herrmann,B., Sorimachi,H. and
Labeit,S.
TITLE The NH2 terminus of titin spans the Z-disc: its interaction with a
novel 19-kD ligand (T-cap) is required for sarcomeric integrity
J. Cell Biol. 143 (4), 1013-1027 (1998)
JOURNAL 93034591
MEDLINE 9817758
PUBMED 9817758
REFERENCE 2 (bases 1 to 6000)
AUTHORS Labeit,S.
TITLE Direct Submission
COMMENT Submitted (16-SEP-1998) S. Labeit, EMBL Heidelberg, Meyerhofstr. 1,
69117 Heidelberg, FRG
FEATURES Related sequences U28657, D83390, X90568.
source Location/Qualifiers
1..6000
/organism="Oryctolagus cuniculus"
/mol_type="mRNA"
/db_xref="taxon:9986"
/chromosome="2"
/map="2q24"
/tissue_type="soleus skeletal muscle"
/dev_stage="adult"
1..6000
/gene="titin"
gene

```

```

CDS
1..>6000
/genes="titin"
/codon_start=1
/product="titin"
/protein_id="CAA77028.1"
/db_xref="GI:3928489"
/db_xref="SPTREMBL:O97791"
/translation="MTTQALFTFQLOPSVVVLEGSAAATFEAHVSGFVPEVSVWRDDQ
VNSTSLPGVQISFSDGRKALTI PAVTKANSGRYSRLATNGSGOATSAELLVTAETA
PNFVORLQSMVROGSOVLQVRVTGIPVVKFYRDLGAEIOSSLOFQISQEGDLVS
LRIAEAYPEDSGTYSVNAIVSGRATSTALLVQGEVWPVKRPTITVSTQPSIAAKQ
TRIEKIEAHFDGARSATVEMVVDGATQOLPHKTPPIIPKPKSRSPTPSPSIAAKQ
LARQSPSPRHSPPSVRVHRAPTSPSVSPAGRISTSPRSVKPSPLLVRKTQTTT
MATGEVPPPKQEGVVASSTAEARETTMTSTTEIRREWEGRYGLQBEVQITIGAA
GAAAVSTSAFAGLGSWFATGAKVQKADKSAVAATVAADVMDVRAVEVIGAVEQ
TAQRTTTTAVHIQPAQEQVRKEAEKIAVSKVVAADAKAEQELKSRREITTTKQEQM
HI THEOLRKETAKAFVPKVVISAAKAKEOETRTITGETTTKQEQKQITQETMMKTRKT
VVPKVIATPKVKEQDLVSRREGISTKREQVOITQEKMRKEAKTALSTIATATAKA
TEQETVLRREGNATQEQIOVTHGVGCKKAEAVATVAADVMDVRAVEVIGAVEQ
DSYAOQTTLIEGYKHEHISATKVAEHPKRPASEPHVVPKAPVVIQAPSETHIKTTEQ
VGMHISQIKKTTDVTSERLVHVDRKPRTPASPHFTVSKISVPKTEHGYEASIAGSAIE
TLQKELSATSSAOKITKSVKAPTVPKAEARVRAETPSPQPFPAATPDYTKSQAGSIEV
KKEGVSIITGSAVREERLEVLHGREAQVEMARVPAEIPVTPPTLVSGLKNMTVIE
GESVTLECHISGYPSTVWYREDYQIESIDQITFQSGIARLMI REAFADRGRT
CSAVNAGTSTSCYLAQVSEFEKETTVAELKTTTEKRFVESRDVMDTITTEEQ
AMPGEAPAFPFTIKPVVQKLVGGSVVFCQIGGNPKPHVYWKSGVPLTTGYRIKVS
YNKOTGECKLIVISMTFADDAGETYIVIRNKHGETSASALLEAEYESLMKSDQEMLY
OTQMTAFVQEPKVEGIAPGVSYSEYEKEQALIRKMAKDTVMVTRTFVEDQEFHI
SSPEERLKEIEVRIIKTLEELLEDGEKMAVDISESAIESGFELRIKNVRILEG
MGVTHFKMNGYPLPKIANYKDGKRIKPGERYOMDFLODGRASIRIPVLPDEGIYT
AFASNIKNAICSGKLYVEPAAPLGAFTYIPTPEVPSIRSVSRVSPIRSPAR
MSPARSPGRLEETDESQERLYKPVFLKPTSFCKLEGO SARFDLKVVGRRPMPET
WFHNGQVQVNDYTHKVIKEDGTQSLIIVPATSDSGEMTVVAQNRAGKSTISVLTV
EAVEHQVKVPVFEKLVNKNVKEGARLEMKVATCNPNPDIVLWKNDSIIVPHKPKII
IEGTGOAALKIDSTVSQDSAWYTATANKAGRTDTRCKVNEVEFAEPEPERLIIIP
RGTYRAKEIAAPALEPLHRLYGOEEDGLYDKQKQKPFKKLSLRLKRFQPAH
PECLRTPIGDPMTVENVLHDGKPLEAANRLRMINEFCYCSLDYGVAVSRDSGVITCRA
TNKIGDTHSPATLIVKDESLVESQIPGKQGLQRIEELERMAHEGALTGVTDQKE
KQKPDIVLPEPVRLEGETARFRVGTYPQPKLWNLNGQLIRKSKRFTVVDGHIH
YLDIVCKSYDTGCVKTAENPEGVIEHKVKLEIQOERDFSVLRRAPEKPEFHIHE
SGKLQFEVQKVDPRVDTSKTEVVKLKRAERI THEKVSESESELSRSPKRRTEBGGYE
AITAVELKSRKDESYEELKTKDELHHTKELTAEKKAL"
BASE COUNT 1859 a 1402 c 1502 g 1236 t 1 others
ORIGIN

Query Match      49.2%; Score 25.6; DB 4; Length 6000;
Best Local Similarity 77.5%; Pred. No. 1.4e+03;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 2 TGCAGTTCTCTCAGGCTCTCATCGCTCTGCTGCTGCGGCGG 41
|||||
Db 2633 TGCAGTACTTCCAGGCGCTCTCGGCGACGCGCTGCGCG 2594
|||||

RESULT 22
AE008035/c
LOCUS      Agrobacterium tumefaciens str. C58 (Cereon)
DEFINITION of 254 of the complete sequence.
ACCESSION AE008035
VERSION AE008035.1
KEYWORDS AE008035.1 GI:15156068
SOURCE Agrobacterium tumefaciens str. C58 (Cereon)
ORGANISM Agrobacterium tumefaciens str. C58 (Cereon)
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
REFERENCE 1 (bases 1 to 10951)
AUTHORS Hinkle,G., Slater,S.C. and Goodner,B.
TITLE Complete Genome Sequence of Agrobacterium tumefaciens C58
(Rhizobium radiobacter C58), the Causative Agent of Crown Gall
Disease in Plants
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 10951)

```

AUTHORS Hinkle,G., Slater,S.C. and Goodner,B.
 TITLE Direct Submission
 JOURNAL Submitted (14-AUG-2001) Cereon Genomics, 45 Sidney Street, Cambridge, MA 02139, USA
 FEATURES
 source Location/Qualifiers
 1.10951
 /organism="Agrobacterium tumefaciens str. C58 (Cereon)"
 /mol_type="genomic DNA"
 /strain="C58"
 /db_xref="taxon:181661"
 complement(97..1656)
 /gene="AGR_C_1928"
 complement(97..1656)
 /gene="AGR_C_1928"
 /notes="(AP001514) BH2174
 unknown conserved protein in B. subtilis"
 /codon_start=1
 /transl_table=11
 /product="AGR_C_1928p"
 /protein_id="AAK86853.1"
 /db_xref="GI:15156071"
 /translation="MNSALCIHHITMITRKVQANVDFYAGFLGLRLVKRTAGFEDAMQ
 LHLVGGDGGSGSLTFLVWGGGGRVCGYGGTLEISVAIDPAAIGFWLTKALKFGL
 RAGPSDEFGAPVIRLKDPLGLVVLKVGTHAFADATPHVTADISADDAICRTYGTATML
 SEVPEETAFLEHFGTRQESQAGAIRRLVSDSGDIIIVRDATGFWSSAPGTGTVDHI
 AFRAKMDDELVTGLKSLNSTTNAHDKYFSLYVREPGGLIEMATDAPGMTVD
 EPTILGCEELFIPPLFMKDEADVVALPQAMPGEERIIVYRDLPFVHRFYTPEEPDGS
 TIILHSGSGSETSLMPLAHRAAPGATLLGVGRSTEEDIARWFRFPDFSPQDKDIL
 LEAFAPAAIVEGAIRAYGLDRSKLRFIGHSGANFYAAFAALYPELAGDALLYRPMVP
 LETWQSDLSGRNPFVLVAGERDKYRGASELQRLTASGARAIEVETVDGHELGLPDL
 EAARRWLSDAAS"
 1948..2991
 /gene="AGR_C_1930"
 1948..2991
 /gene="AGR_C_1930"
 /notes="hypothetical protein"
 /codon_start=1
 /transl_table=11
 /product="AGR_C_1930p"
 /protein_id="AAK86854.1"
 /db_xref="GI:15156070"
 /translation="MLTRLLRSLTALIAIGAVAGFSLPATGAYAEPTVPKFTLDKWF
 EGPAAGFLALDGGYKFAEGLDVTIISGNSVEAIPRVATGAYQMGFGDINSYMKFLD
 EDPSQIKAVMYVERPTFAVGRKSLGVTGDPKLEGGKLGAPPPDGAQWPAFKQ
 VAGLDDSKIKIESGFVREPMLAKGVDVAFGFAFSLVILNKKGIADDDISTILMA
 EHLGNLYGNALVNTDFAEKNPEAVKGFLLAKAKGADSVKTPPEEGVAAVLKRNETLD
 SATELERLANMANSMTIKTPYVVENGGVDDPARLAASLDTLKVSMGLKGNVKAQVFD
 AAVLPKESRMLP"
 3242..4030
 /gene="AGR_C_1932"
 3242..4030
 /gene="AGR_C_1932"
 /notes="probable ATP-binding component of ABC transporter
 Pa2294 [imported] - Pseudomonas aeruginosa (strain PAO1)"
 /codon_start=1
 /transl_table=11
 /product="AGR_C_1932p"
 /protein_id="AAK86855.1"
 /db_xref="GI:15156071"
 /translation="MSHLVEIDSVMRYGSGACTLAVSGLNLAVDKGEFAAAVVGPSGC
 GKSLMKLVTLGHIPIRGNVIVAGRQVTEPVSIVGNFQNTMLPWRITLLENILLPLE
 IVEKRRRURAHKASYAKAERLLEIVLGLKFGSKYPQLSGGMOQRANLCRALIHPE
 ELMLDPEFGALDAFTREELWCVIRDHAAQRVTIILVTHDLREATFLADKIFVMSAR
 PGRVLAHRVPFARPRQLDIMYDKGFNDVMVQLHGHIAQARVAA"
 4027..4845
 /gene="AGR_C_1933"
 4027..4845
 /gene="AGR_C_1933"
 /notes="probable permease of ABC transporter PA3512
 (imported) - Pseudomonas aeruginosa (strain PAO1)"
 /codon_start=1
 /transl_table=11
 /product="AGR_C_1933p"

/protein_id="AAK86856.1"
 /db_xref="GI:15156072"
 /translation="MSAINESSVAAPRKPLIERVNVVKAAPWLYTLALFALWELLYV
 ALKMPPTILSPIRVQVIVQVWSPVWKNLSQTLDTYTLTGFAIAAAGLAGLGFIGWS
 KTIYAGLYPLMIGFNAIPVILVIFWFGIGTVPVAVLTAFLISFFPIVNVATGLA
 TIETPEDVLRALGAKKMDIMLKVGIPRSMYPFGSLKIAITLAFVGSVSVSETVASNY
 GLGNMSSAQSQFNVLVFAGLLMLAVEGIAVMVAMWALEKRMKTGMWHRSTMQ"
 complement(4897..5580)
 /gene="AGR_C_1935"
 complement(4897..5580)
 /gene="AGR_C_1935"
 /notes="HYPOTHEICAL 28.9 KD PROTEIN IN AADK-SIGZ
 INTERGENIC REGION"
 /codon_start=1
 /transl_table=11
 /product="AGR_C_1935p"
 /protein_id="AAK86857.1"
 /db_xref="GI:15156073"
 /translation="MOKLTUVTGALALGSMVAAAAATAAGDGKPAASHNHGHGHS
 QMTEAEKQIYGYFEDAQVKPRPLTDWEGDMQSVYPYLLNGSLDSVNAEKAAGDKTA
 ABEYAEYIGYKTDVNRIVIAPSVTFYKEGKTGTYVSDGQELITYKKGNRGVYII
 FKKTAGDEAAPQIQISDHTIAPKAGHYHLYWGNDRKALMDVEVTNMTPTTYPFLMDGA
 DIVENAAH"
 complement(5758..6579)
 /gene="AGR_C_1937"
 complement(5758..6579)
 /gene="AGR_C_1937"
 /notes="HYPOTHEICAL 30.9 KDA PROTEIN IN SBM-FBA INTERGENIC
 REGION (ORF 4)"
 /codon_start=1
 /transl_table=11
 /product="AGR_C_1937p"
 /protein_id="AAK86858.1"
 /db_xref="GI:15156074"
 /translation="MEQQTADILVASOTAVROASALAVQVSFVIGALLLLIIGWLAA
 AFLQWAFQGLSRIRGFDATLAGFAGGIRYVVLIVVNLQFGQVQTASILAALGA
 AGLAIGLQGLTQNIAGMLLVLPFRFVGEIETVTSVGTIIEGLFATELKTSOG
 LYRLAPNSTLMNVPIITNYSRFPSSRRHLSLTVKNDEDMAAQDMLMTVVSSESRLLD
 PAPVTFVDSVTADSATKLYRWVRSNDYFTTDDVTDKMRLAFDERKAEVNAQPA"
 complement(6657..7175)
 /gene="AGR_C_1940"
 complement(6657..7175)
 /gene="AGR_C_1940"
 /notes="hypothetical protein"
 /codon_start=1
 /transl_table=11
 /product="AGR_C_1940p"
 /protein_id="AAK86859.1"
 /db_xref="GI:15156075"
 /translation="MTCIVVPSLSRIAEMAVKHKAREMVTLIAKEOSFHRPAPVAAER
 HLTLMNDIAFYKCTGLIAPDDAHVLKLIDFAGENQOSAPLLIHCWNGISRSPPAAAVI
 AALLYPPDQDETELAQRLRTVSPYTPNSRIIAIGDRLLGREGRLVKAIIKAIGRGADT
 DGNVPVPLPLAG"
 complement(7172..7852)
 /gene="AGR_C_1942"
 complement(7172..7852)
 /gene="AGR_C_1942"
 /notes="HYPOTHEICAL 21.6 KD PROTEIN IN AHCY-HVRRB
 INTERGENIC REGION (ORF5)"
 /codon_start=1
 /transl_table=11
 /product="AGR_C_1942p"
 /protein_id="AAK86860.1"
 /db_xref="GI:15156076"
 /translation="MERAFIPGRRSGSGGLRVAPAKSPRAWORMLSGRRLLDLPDS
 LHDVEIADIAHGLARVWNGQTRGDHAFVTAQCHLIVETIFCRMCPGTPDPMQNAL
 LHDPEYVIGDMISPFKSVVGGYKTVKEKRLEAAVHRLFGLPHPHAKRELKDKRIKKADT
 VAAFEATELAGFSTAEAKQKFFGLPRGITRDMFDIIPLPSTEAQRLFIARFEATLRL
 VTRTGAV"
 complement(7798..8646)
 /gene="AGR_C_1943"
 complement(7798..8646)
 /gene="AGR_C_1943"

/note="hypothetical protein RP464 - Rickettsia prowazekii"
/codon_start=1
/transl_table=1
/product="AGR_C 1943p"
/protein_id="AAK86861.1"
/db_xref="GI:15156077"
/translation="MSSAFADRLRLIRVSGTGAEEFLNNLTADIENTPEGETRASAL
LTPQKILFDLIRWDGDRDYLVETGAEEQDALLRRLTYKLRARVELKAEVTEIGVGF
WGNVSTGAVRGDFGAFKAGVDPVPGASGEAAAEALRVEHGIAESGRDYALQDA
FPFDLMDVNDGVSKGCFVGVQVWRMKHGTARRRVTVTSADGTLPASGTEITAN
GKPGVAGLTGVGNALATVTRDVRADALAAGTPLLADNVAVSVAPWSGLSFPAADP
AARAED"
complement(8769..9536)
/gene="AGR_C 1946"
complement(8769..9536)
/gene="AGR_C 1946"
/note="hypothetical protein"
/codon_start=1
/transl_table=1
/product="AGR_C 1946p"
/protein_id="AAK86862.1"
/db_xref="GI:15156078"
/translation="MCLSKGLRSSLILLSLALTAGPIGLPEKAFALSCLKPAQKG
ESEPAQNGEQOQAPAEDEVEGSDGILPDPPLVDRSAGAKQOSAKPDTTPEMS
WIEHDSKAPVPRKLRQOIIDAASGDIKLPFPVAGONOFRIDGSEDPIAL
KSYSGDPDGLVLAIIIDLLSTGYAHIDAGTPDEAYVFPYPAGKPLNTLTAPEKVELL
RIITAGDADMGEYGNYSFYRIGISPDGKWKFXTAGD"
complement(9551..9808)
/gene="AGR_C 1947"
complement(9551..9808)
/gene="AGR_C 1947"
/note="hypothetical protein"
/codon_start=1
/transl_table=1
/product="AGR_C 1947p"
/protein_id="AAK86863.1"
/db_xref="GI:15156079"
Query Match 49.2%; Score 25.6; DB 1; Length 10951;
Best Local Similarity 70.8%; Pred. No. 1.3e+03;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Oy 2 TGCAGTCTCTCAGCGCTCTCATCGCTCTCGTGGCTGCGCGCTCGCCCA 49
|||||
Db 6308 TGCAGACGCGCTCGATTCTCGCCGCTCGCGCTGCGCGCTCGCCCA 6261
RESULT 23
AE009069/c
LOCUS
DEFINITION Agrobacterium tumefaciens strain C58 linear BCT 20-DEC-2001
95 of 256 of the complete sequence.
ACCESSION AE009069.1 GI:17739437
VERSION
KEYWORDS
SOURCE
ORGANISM Agrobacterium tumefaciens str. C58 (U. Washington)
Agrobacterium tumefaciens str. C58 (U. Washington)
Bacteria: Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
1 (bases 1 to 11764)
Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.,
Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F.,
Zhou, Y., Bovee Sr., D., Chapman, P., Clendenning, J., Deatherage, G.,
Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M.,
McClelland, E., Palmeri, A., Raymond, C., Rouse, G.,
Saenphimachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I.,
Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M.,
Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V.
and Nester, E.W.
The genome of the natural genetic engineer Agrobacterium
tumefaciens C58
Science 294 (5550), 2317-2323 (2001)

MEDLINE
PUBMED
REFERENCE
AUTHORS

21608550
11743193
2 (bases 1 to 11764)
Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.,
Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F.,
Zhou, Y., Bovee Sr., D., Chapman, P., Clendenning, J., Deatherage, G.,
Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M.,
McClelland, E., Palmeri, A., Raymond, C., Rouse, G.,
Saenphimachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I.,
Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M.,
Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V.
and Nester, E.W.
Direct Submission
Submitted (27-SEP-2001) Department of Microbiology, University of
Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA
98195-7242, USA
Location/Qualifiers
1..11764
/organism="Agrobacterium tumefaciens str. C58 (U.
Washington)"
/mol_type="genomic DNA"
/strain="C58"
/db_xref="taxon:180835"
177..1220
/gene="Atul046"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="ABC transporter, substrate binding protein"
/protein_id="AAL42059.1"
/db_xref="GI:17739438"
/translation="MILTRLLSRTLAIAIGAVGSLPATGAYETPVKFTLDWKFLD
EGPAGELLALDKGYFKAEGLDVTIDSGNSVEALPRVATGAYOMGFGDINSVMKFLD
EDSQKIKAVMYETPTFAVGRKSGVDGPKSLGKGLGAPPGDAPQAFQWPAFKO
VAGLDGFKIKIESIGFPPVPMKAGSDVAVGFAFVILNLKKGITADDDISTILMA
EHLNLGVAVLVNTDFAEKNPFAVKGLAKALAKGADSVKTPPEGVAAVLKRNELTD
SALELRNLNMANSMNIKTPYVVENGMGVDPARLAASLDTLKVSMLKGNVKAEOVFD
AAYLPKPEERMLP"
1471..2259
/gene="Atul047"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="ABC transporter, nucleotide binding/ATPase
protein"
/protein_id="AAL42060.1"
/db_xref="GI:17739439"
/translation="MSHLVIDSVDMRYGSGAGTLAVSGLNLAVDKCEFAAVGPGSGC
GKSTLMKLVTLGHIPOQRNVIVAGRQVTEPVSIVGMAFQNTMLPWRITLNLILLPLE
IVEKRRRLRAHKASYAKAELELLEIVLGLGKSGYKQJSGGQOORALNCRLAIHEP
ELMLDPEFGALDAFTREELWCVRIDJHAQRVTIILVTHDLREATFLADKIFVMSAR
PGRVLAERVPFFARPQLDIMYDKGFNDVMQELHGEIAQARVAA"
2256..3074
/gene="Atul048"
2256..3074
/gene="Atul048"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="ABC transporter, membrane spanning protein"
/protein_id="AAL42061.1"
/db_xref="GI:17739440"
/translation="MSAIMESSVAAEPRKPLIERVNVNKAAPMLYTLALFALWELLVY
ALKMPPTILPSPIRVGOVIVQYWSPIWNSLQTLYTTTLGFAIYAAAGLAIGLFIGWS
KTIYAGLYPLMIGFNAIPKVALPILVIFGIGTVPALVATFLISFFPIVNVATGLA

```

TTPETEDVLRLGAKKMDIMLVKGIPIRSMPIYFFGSLKIAITLAFVGSVVSETVASNY
GLGNMSSAQSQFNPLVFLVAGLLMLAVEGIAMVAVMAWLEKRMGTGWAHRSTMGQ"
/genes="Atul049"
complement(3126. .3809)
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL42062.1"
/db_xref="GI:17739441"
/translacion="MQKTLTRVTCALALGMLVAAAAATAAGCKKPAASHNHGHDS
QMTAEKQYKGFEDQVQKPRPLTDMEGDMQSVYPYLLNGSLDSVMAEKAAGDKTA
AEYRAYEIGYKTDVNRIVIAQDSVTFYKEGKPTGYVSDGOEILTYKGNRGVYI
FKKTACDEAAQPIQFSDHTIAPKAGHYLYWGNDRKALMDEVNTWPTYPLMDGA
DIVEENAAH"
complement(3987. .4808)
/genes="Atul050"
complement(3987. .4808)
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL42063.1"
/db_xref="GI:17739442"
/translacion="MEQQATDILVASQTVARQASALAVQVSFVIGALLLLIIGWLAA
ALQRAWFQGLSRIRGDFAGGIRYVVLIVVILVQFGVQVQASILAALGA
AGLALGALOGTLONIAAGMLLVLPFRVGEIYTSVGTIIEIGLFATELKTSDG
LYRLAPNSTLMNPVITNYSRFRSRRLHSILTVKNDDEMAAODMLTVVRSRILLD
PAPVTFVSDVTSATVTKLRYWVRSDNYFVTTDRVTKAMRLAFDERKAEVNAQPA"
complement(4886. .5404)
/genes="Atul051"
complement(4886. .5404)
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL42064.1"
/db_xref="GI:17739443"
/translacion="MTCIVVSPSLRIAEMAVKHKAREMVTLIAKEQSFHRPAVAAER
HUTLAMDIAFKGTGLIAPDDAHVLKLIIDFAGEWDQSAPLLHCWMGISRSPAARI
AALSVPDODETELAQRLRTVSPVTPNSRIIAIGRLDREGRLVKAIKATGRGADT
DGNVPVPLAC"
complement(5401. .6024)
/genes="Atul052"
complement(5401. .6024)
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL42065.1"
/db_xref="GI:17739444"
/translacion="WAPAKSPRAWOMLSGRLLDLPSPLOVETADIAHGLARVARW
NGQTRGDHAFYVAQCLIVETIFCRNCPGATFDEMGMALHDPAPEYVIGDMISPKSV
VGGYKTVKRELEAAVHLRFLGPPHASPARELKRIFKADTVAAAFATELAGFSTAEAQ
KFFGLPRGTRIDMFIDILPLPSTEAQRLFIARFEATETLRVTRTGAV"
complement(6027. .6875)
/genes="gcvt"
/notes="synonym: Atul053"
complement(6027. .6875)
/genes="gcvt"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11

```

```

/product="glycine cleavage system T protein,
aminomethyltransferase"
/protein_id="AAL42066.1"
/db_xref="GI:17739445"
/translacion="MSSAFLLDRRLRVSGTGAEEFLNNLITADIEHNLPGETRASAL
LTPQKILFLFDIRDRDYLVETGAEEQDALLRLMYLRAPVLEKAEVTEGIVGF
WNSVTVTEAGVRDGRFAKAGVRRVPGASASGEAAAYEARLVHGEIASEGRDIALOA
PHDVLMDVNDGVSKKCGFVGOEYVRMKHGRGTARRRVTVVSADGTLPASGTEITAN
GKPVGALVTYGNRALAIVTRDRVADALAAGTPLLADNVAVSVALPAWSGLSFFAADP
AARAE"
complement(6998. .7765)
/genes="Atul054"
complement(6998. .7765)
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL42067.1"
/db_xref="GI:17739446"
/translacion="MSCLSKGLLRSSALLSILLALTAGPIGLPEKAFALSELKPAQKG
ESEPQNGEQEQQAQPAEPDEVEDSGDGLPDLVDRSAGKANQOSAKPDTTPEMS
VIEHDISKAPEVRKLQOIIDAAASGDIEKLRFVIVAGQNFRIQSDSDSEDFIAL
KSYSGDPDGLVLAIIIDLLSTGYAHIDAGTPEAYVFPYFAGKPLNTLTAPEKVELL
RIITAGDLADMOEYGNYSFVRIGISPDGKWKFFTAGD"
complement(7780. .8037)
/genes="Atul055"
complement(7780. .8037)
/genes="Atul055"

Query Match 49.2%; Score 25.6; DB 1; Length 11764;
Best Local Similarity 70.8%; Pred. No. 1.3e+03;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 2 TGCAGTTCTCTACGCTCTCATGCTCTCGTCGCTCGCGCCTCGCCA 49
||||| |
Db 4537 TGCAGCGCGCTCGATTCTCGCGCCTCGCGCCTCGCGCCTCGCCA 4490

RESULT 24
AP005454
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,
PAC clone: P0455F03, complete sequence.
ACCESSION
AP005454
VERSION
AP005454.2 GI:29824104
KEYWORDS
HTG.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1
Sasaki,T., Matsumoto,T. and Katayose,Y.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
clone:P0455F03
JOURNAL
Published Only in Database (2002)
AUTHORS
Sasaki,T., Matsumoto,T. and Katayose,Y.
TITLE
Direct Submission
JOURNAL
Submitted (20-JUN-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Apr 11, 2003 this sequence version replaced gi:21624006.
The orientation of the sequence is from SP6 to T7 of the PAC clone.
COMMENT
Location/Qualifiers
1 .106732
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"

```

```

/db xref="taxon:39947"
/chromosomes="7"
/clones="P0455F03"
BASE COUNT 30378 a 23692 c 22863 g 29799 t
ORIGIN

Query Match 49.2%; Score 25.6; DB 8; Length 106732;
Best Local Similarity 70.8%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 ATGCAGTCTCTCAGCTCTCATCGCTCTCGTCCGCTCGCGCTCGCC 48
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 50531 ATGCAGTCTCTTCCGCGCTCGCGCTCGCGCTCGCGCTCGCC 50578

RESULT 25
AP003513/c
LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 6 clone P0690H04,
*** SEQUENCING IN PROGRESS ***.
ACCESSION AP003513
VERSION AP003513.1 GI:13603473
KEYWORDS HTG; HTGS PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
clone: P0690H04
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 139487)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (11-APR-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
COMMENT NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1..139487
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosomes="6"
/clones="P0690H04"
BASE COUNT 40074 a 30224 c 29970 g 39019 t 200 others
ORIGIN

Query Match 49.2%; Score 25.6; DB 2; Length 139487;
Best Local Similarity 77.5%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 9 CTCTCAGCTCTCATCGCTCTCGTCCGCTCGCGCTCGCC 48
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101198 CGTCTATGCTCTCATCGCTCTCGTCCGCTCGCGCTCGCC 101159

RESULT 26
AP003569/c
LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 6 clone P0425F05,
*** SEQUENCING IN PROGRESS ***.
ACCESSION AP003569
VERSION AP003569.1 GI:13936409
KEYWORDS HTG; HTGS PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
clone: P0425F05
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 164978)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
COMMENT NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1..164978
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosomes="6"
/clones="P0425F05"
BASE COUNT 47386 a 34709 c 34874 g 47659 t 350 others
ORIGIN

Query Match 49.2%; Score 25.6; DB 2; Length 164978;
Best Local Similarity 77.5%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 9 CTCTCAGCTCTCATCGCTCTCGTCCGCTCGCGCTCGCC 48
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 17406 CGTCTATGCTCTCATCGCTCTCGTCCGCTCGCGCTCGCC 17367

RESULT 27
AC114896
LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 3 clone
OSJNBa0091B22, *** SEQUENCING IN PROGRESS ***. 5 ordered pieces.
ACCESSION AC114896
VERSION AC114896.3 GI:20503086
KEYWORDS HTG; HTGS PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Tallon, L., Gansberger, K.,
Kim, M., Overton II, L., Bera, J., Tsitrin, T., Krol, M., Jarrahi, B.,
Jin, S., Koo, H., Zismann, V., Hsiao, J., Blunt, S., Vanaken, S.,
Utterbach, T., Feldblyum, T., Yang, Q., Haas, B., Suh, B., Peterson, J.,
Quackenbush, J., White, O., Salzberg, S. and Fraser, C.

```

```

LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 6 clone P0425F05,
*** SEQUENCING IN PROGRESS ***.
ACCESSION AP003569
VERSION AP003569.1 GI:13936409
KEYWORDS HTG; HTGS PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
clone: P0425F05
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 164978)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
COMMENT NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1..164978
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosomes="6"
/clones="P0425F05"
BASE COUNT 47386 a 34709 c 34874 g 47659 t 350 others
ORIGIN

Query Match 49.2%; Score 25.6; DB 2; Length 164978;
Best Local Similarity 77.5%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 9 CTCTCAGCTCTCATCGCTCTCGTCCGCTCGCGCTCGCC 48
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 17406 CGTCTATGCTCTCATCGCTCTCGTCCGCTCGCGCTCGCC 17367

RESULT 27
AC114896
LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 3 clone
OSJNBa0091B22, *** SEQUENCING IN PROGRESS ***. 5 ordered pieces.
ACCESSION AC114896
VERSION AC114896.3 GI:20503086
KEYWORDS HTG; HTGS PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Tallon, L., Gansberger, K.,
Kim, M., Overton II, L., Bera, J., Tsitrin, T., Krol, M., Jarrahi, B.,
Jin, S., Koo, H., Zismann, V., Hsiao, J., Blunt, S., Vanaken, S.,
Utterbach, T., Feldblyum, T., Yang, Q., Haas, B., Suh, B., Peterson, J.,
Quackenbush, J., White, O., Salzberg, S. and Fraser, C.

```


AUTHORS TITLE JOURNAL

Buell,R.
Direct Submission
Submitted (08-MAY-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
On May 8, 2002 this sequence version replaced gi:20146765.

COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 21285: contig of 21285 bp in length
* 21286 21385: gap of unknown length
* 21386 59264: contig of 36879 bp in length
* 58265 59364: gap of unknown length
* 58355 71533: contig of 13169 bp in length
* 71534 71633: gap of unknown length
* 71634 73850: contig of 2217 bp in length
* 73851 82966: gap of unknown length
* 82967 83066: contig of 9016 bp in length
* 83067 118460: contig of 35394 bp in length
* 118461 118560: gap of unknown length
* 118561 144077: contig of 25517 bp in length.

FEATURES source

1. 144077
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OSJNBA0036E17"
/note="japonica cultivar-group"

BASE COUNT 41309 a 31223 C 30860 G 40084 t 601 others
ORIGIN

Query Match 48.8%; Score 25.4; DB 2; Length 144077;
Best Local Similarity 68.6%; Pred. No. 1.2e+03;
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 2 TGCAGTCTCTCAGCTCTCATCGCTCTCGTTCGCGCGCTCGCCAGTG 52

Db 135221 TGCAGTCTCTCAGCTCTCTTGTCTCAGCTCGCAGCGTCTCGCTCTTG 135171

RESULT 30 CNS08C87/c

LOCUS CNS08C87 152359 bp DNA linear HTG 07-MAR-2003
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 12 clone
OSJNBA0029K06, *** SEQUENCING IN PROGRESS ***.

ACCESSION

AL731758

VERSION AL731758.3 GI:28892654

KEYWORDS HTG; HTGS PHASE2; HTGS ACTIVEPIN.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

1 (bases 1 to 152359)

REFERENCE
AUTHORS Choisine,N., Orjeda,G., Cattolico,L., Demange,N., Wincker,P.,
Seguens,B., Pelletier,E., Scarpelli,C., Salanoubat,M.,
Weissenbach,J. and Quetier,F.

TITLE Oryza sativa chromosome 12 sequencing

REFERENCE 2 (bases 1 to 152359)

JOURNAL AUTHORS

Unpublished
Genoscope.
Direct Submission
Submitted (06-MAR-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

On Mar 9, 2003 this sequence version replaced gi:28460535.
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: <http://www.genoscope.cns.fr/>
Contact: SegRef@genoscope.cns.fr

The following sequence is oriented from the T7 to the SP6 end.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence. Work on the sequence is in progress
and the release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated with foreign sequence from E.coli, yeast, vector,
phage, etc. . . even if efforts are made to eliminate these
contaminating sequences.

The nucleotide sequence of this BAC clone was generated by
combining Syngenta and Genoscope sequencing data.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 152359: contig of 152359 bp in length.

FEATURES source

1. 152359
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="12"
/clone="OSJNBA0029K06"
/clone_lib="OSJNBA"

BASE COUNT 43054 a 34046 C 33464 G 41741 t 54 others
ORIGIN

Query Match 48.8%; Score 25.4; DB 2; Length 152359;
Best Local Similarity 68.6%; Pred. No. 1.2e+03;
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 2 TGCAGTCTCTCAGCTCTCATCGCTCTCGTTCGCGCGCTCGCCAGTG 52

Db 110057 TGCAGCCGCTCGCGTCTTGTCTCTCGCCGAGCGTCTTCTCTTG 110007

RESULT 31 CNS08CAF/c

LOCUS CNS08CAF 168864 bp DNA linear HTG 07-MAR-2003
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 12 clone
OSJNBA0083D13, *** SEQUENCING IN PROGRESS ***.

ACCESSION

AL831802

VERSION AL831802.4 GI:28892660

KEYWORDS HTG; HTGS PHASE2; HTGS ACTIVEPIN.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

1 (bases 1 to 168864)

REFERENCE
AUTHORS Choisine,N., Orjeda,G., Cattolico,L., Demange,N., Wincker,P.,
Seguens,B., Pelletier,E., Scarpelli,C., Salanoubat,M.,
Weissenbach,J. and Quetier,F.

TITLE Oryza sativa chromosome 12 sequencing

REFERENCE 2 (bases 1 to 168864)

JOURNAL AUTHORS

Unpublished
Genoscope.
Direct Submission
Submitted (06-MAR-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

On Mar 9, 2003 this sequence version replaced gi:28460587.
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr

The following sequence is oriented from the T7 to the SP6 end.
IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continue. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage, etc. . even if efforts are made to eliminate these contaminating sequences.

The nucleotide sequence of this BAC clone was generated by combining Syngenta and Genoscope sequencing data.

Contigs composition :

166717 bp contig from 1 to 166717
2097 bp contig from 166768 to 168864.

- * NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs
- * are represented as runs of N. The order of the pieces
- * is believed to be correct as given, however the sizes
- * of the gaps between them are based on estimates that have
- * provided by the submittor.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

- * 1 168864: contig of 168864 bp in length.

FEATURES

source	Location/Qualifiers
LOCUS	1..168864
	/organism="Oryza sativa (japonica cultivar-group)"
	/mol_type="genomic DNA"
	/cultivar="Nipponbare"
	/sub_species="japonica"
	/db_xref="taxon:39947"
	/chromosome="12"
	/clone="OSJNBa0083D13"
	/clone_lib="OSJNBa"

BASE COUNT	45973 a 38589 c 38829 g 45421 t 52 others
ORIGIN	

Query Match	48.8%; Score 25.4; DB 2; Length 168864;
Best Local Similarity	68.6%; Pred. No. 1.2e+03;
Matches	35; Conservative 0; Mismatches 16; Indels 0; Gaps 3

Qy	2	TGCAGTCTCTCAGGCTTCATCGCTCTCGCTGCCGGCCCTGCCAGTG	52
Dd	7885	TGACGCCCGCTCGGTTCTTGTTCCTCGCCGACCGTCTTGTCTTG	7835

RESULT 32

SC0939129/c

LOCUS SC0939129 292200 bp DNA linear BCT 11-FEB-2003

Streptomyces coelicolor A3(2) complete genome; segment 26/29.

AL93129 AL049485 AL049587 AL079355 AL138598 AL138667 AL138668

AL138977 AL353864 AL353870 AL356593 AL356812 AL645882

AL939129.1 GI:24418971

. Streptomyces coelicolor A3(2)

Streptomyces coelicolor A3(2)

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

1 Bentley,S.D., Chater,K.F., Carrdeno-Tarraga,A.M., Challis,G.L., Thomson,N.R., James,K.D., Harris,D.E., Quail,M.A., Kieser,H., Harper,D., Bateman,A., Brown,S., Chandra.G., Chen,C.W., Collins,M., Cronin,A., Fraser,A., Goble,A., Hidalgo,J., Hornsby,T., Howarth,S., Huang,C.H., Kieser,T., Lark,L., Murphy,L., Oliver,K., O'Neill,S., Rabinowitz,E., Rajandream,M.A., Rutherford,K., Rutter,S., Seeger,K., Saunders,D., Sharp,S., Squares,R., Squares,S., Taylor,K., Warren,T., Wietzorek,A., Woodward,J., Barrall,B.G., Parkhill,J. and Hopwood,D.A.

```

identity in 342 aa overlap). Contains Pfam matches to
entry PF00532 Peripla_BP_like. Periplasmic binding
proteins and Laci family and to entry PF00356 laciI,
Bacterial regulatory proteins, laci family. Contains
probable helix-turn-helix motif at aa 9-30 (Score 2307,
+7.04 SD)
/codon_start=1
/transl_table=11
/product="putative transcriptional regulator"
/protein_id="CAB45587.1"
/db_xref="GI:5139569"
/translations="MTETGPRPTLEAVERAGVSRTSVRVNGHGVDRDALAERVRR
AVELEGVPHNHAASLVTRRHOAVVAAEPETRVFADPVFAOQLAGISKELTAHDNO
LVLLLEGEDHARVGLYLAGHVGDGALVESLHDLPLGLVRSAGVPTVPGRPDWD
DGRDDVVYDSDNRGNARSVHHLGLGRTRIAHTIGPLDQTSAAKLAFGRVRFCA
GPGVLARGDFTSGGGERAMRELDRCPLDAVFAANDLTAAAGLRLVLRGRGRVDDV
AVVGFDDMLPVAEQTPPLTTVTRQDIEGMGLMARLLRLGLDRRAADATDFAARTPT
AAPGGVVLPTTLVHRSTA"
1291..1374
/gene="SC06713"
/notes="Pfam match to entry PF00356 laciI, Bacterial
regulatory proteins, laci family, score 39.70, E-value
9.4e-10"
1465..2202
/gene="SC06713"
/notes="Pfam match to entry PF00532 Peripla_BP_like,
periplasmic binding proteins and Laci family., score
95.60, E-value 9.5e-25"
/complement(2088..2111)
/notes="repeat 7; 23 bp imperfect inverted repeat"
/gene="SC06714"
/notes="synonym: SC4C6.24c"
/complement(2440..3231)
/gene="SC06714"
/notes="SC4C6.24c, possible hydroxylase, len: 263 aa;
similar to TR:Q55078 (EMBL:U50973) Streptomyces sp. strain
C5 daunomycin C-14 hydroxylase (275 aa), fasta scores:
opt: 422 z-score: 482.3 E(): 1.5e-19, 38.5% identity in
273 aa overlap. Similar to TR:Q50527 (EMBL:AL009204)
S.coelicolor hypothetical protein (267 aa) (37.9% identity
in 253 aa overlap) and to (EMBL:AL049863), SC5H1.09c.
S.coelicolor possible hydroxylase (265 aa) (35.7% identity
in 258 aa overlap). Also similar to hypothetical proteins
from Mycobacterium tuberculosis and to Mycobacterium
tuberculosis 27.3 kDa MAB HBT7 reactive antigen
(EMBL:AJ007737) (260 aa) (36.7% identity in 264 aa
overlap)"
/codon_start=1
/transl_table=11
/product="putative hydroxylase"
/protein_id="CAB45588.1"
/db_xref="GI:5139570"
/translations="MLTTRVTVGAPNMLDVGTSDLDGATSYGGLFGWRFRSAGPEAG
GYAFFELDRIVAGMTTEEQPPSWTVYFOAPDARAAQAAGHGVLVGPMDVM
DOGTMAILADRALGPGIWPQGRAGLDVGTGALCOWELHTADIAAAAYYRAVLG
LETSGVSPFGSVCVNPAGEGDMFGGLVPLAEDPADTDADAGALPYFAVDDAOTA
VARTELGGTVRPATDIEGVRVARLDYPGARFAVLRPAPRQG"
/complement(3239..3244)
/notes="possible RBS"
3425..3670
/gene="SC06715"
/notes="synonym: SC4C6.25"
3425..3670
/gene="SC06715"
/notes="SC4C6.25, probable transcriptional regulator, len:
81 aa, similar to many transcriptional regulators (some
putative) from Streptomyces e.g. TR:Q53963 (EMBL:X62287),
WhiB, S.coelicolor transcriptional regulator essential for
sporulation (87 aa), fasta scores, opt: 336 z-score: 452.7
E(): 5.7e-18, 61.1% identity in 72 aa overlap. Also
similar to e.g. TR:Q69649 (EMBL:AL02212) Mycobacterium
tuberculosis putative regulatory protein (100 aa) (40.0%

```

```

identity in 75 aa overlap)"
/codon_start=1
/transl_table=11
/product="putative transcriptional regulator"
/protein_id="CAB45589.1"
/db_xref="GI:5139571"
/translations="MHTETTAAPDLSNQEBALCAQTGGDFFPFPGSSVDRAKRICAL
CPIRSTCLEFALSNDERFGVMGSLSEKRLALRTTP"
4003..4197
/gene="SC06716"
/notes="synonym: SC4C6.26"
4003..4197
/gene="SC06716"
/notes="SC4C6.26, hypothetical protein, len: 64 aa; unknown
function, probable CDS suggested by positional base
preference, GC frame plot and amino acid composition"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB45590.1"
/db_xref="GI:5139572"
/translations="MCRAGGVAALLTEAAABPLAVTGALGCECDRPAPWRELPRAP
RRHPAPGVRAAPEITAHE"
/complement(4242..5228)
/gene="SC06717"
/notes="synonym: SC4C6.27c"
/complement(4242..5228)
/gene="SC06717"
/notes="SC4C6.27c, probable acyl-(acyl-carrier protein)
desaturase, len: 328 aa; similar to many e.g.
SM-STAD LINUS (EMBL:X70962) Linum usitatissimum
stearoyl-ACP desaturase (396 aa), fasta scores; opt: 385
z-score: 461.5 E(): 2.2e-18, 30.2% identity in 311 aa
overlap. The N-terminal 100 aa are divergent. Highest
Query Match 48.8% Score 25.4; DB 1; Length 292200;
Best Local Similarity 68.8%; Pred No. 1.1e+03;
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Oy 1 ATCGAGTTCTCACGCTCTCATCGCTCTCGTCGTCGCGGCGCTCGGCAGT 51
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279401 ATCGGTACCAACGCGCGCGCTCGTCGCTCGCTCGCTCGTGT 279351

RESULT 33
AE005738/c 10037 bp DNA linear BCT 12-JUN-2002
LOCUS Caulobacter crescentus CB15 section 64 of 359 of the complete
DEFINITION genome.
ACCESSION AE005738
VERSION AE005738.1 GI:13421828
KEYWORDS CAULOBACTER CRESSCENTUS CB15
SOURCE Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
ORGANISM Caulobacteraceae; Caulobacter.
REFERENCE 1 (bases 1 to 10037)
AUTHORS Nierman,W.C., Feldblyum,T.V., Laub,M.T., Paulsen,I.T., Nelson,K.E.,
Eisen,J., Heidelberg,J.F., Alley,M.R., Ohta,N., Maddock,J.R.,
Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D.,
Ely,B., DeBoy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L.,
Haft,D.H., Kolonay,J.F., Smit,J., Craven,M.B., Khouri,H.,
Shetty,J., Berry,K., Ufferback,T., Tran,K., Wolf,A., Vamathevan,J.,
Ermoiaeva,M., White,O., Salzberg,S.L., Venter,J.C., Shapiro,L. and
Fraser,C.M.
TITLE Complete genome sequence of Caulobacter crescentus
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)
MEDLINE 21173698
PUBMED 11259647
REFERENCE 2 (bases 1 to 10037)
AUTHORS Nierman,W.C., Feldblyum,T.V., Paulsen,I.T., Nelson,K.E., Eisen,J.,
Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,J.R., Potocka,I.,
Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B.,

```

TITLE
JOURNAL
FEATURES
source
1. 10037
/organism="Caulobacter crescentus CB15"
/mol_type="genomic DNA"
/strain="CB15"
/db_xref="taxon:190650"
224. .2059
/gene="CC0620"
224. .2059
/note="CC0620"
/codon_start=1
/transl_table=1
/product="type I restriction-modification system, M subunit, putative"
/protein_id="AAK22605.1"
/db_xref="GI:13421829"
/translation="MNHOSLSAFIWSVADLLRGDYKSDYGVKILPFTVLRRLDCVLE
PTKAAVLAEHDKRAAGVDPFPLRVGAGFYNTKQMDLTVLGDADNVAANL.YAVL
QATSPVARDI.FERPEFHAQIERLAKGLLVMAEKFTRIDLHPEAVNHONGHVFEE
IRKFAELSNAGTGHFTPREVIRLMVELI.FVEDDAALSEGVVRTIYDPTAGTGMLS
VAERLLNQTPGAKSYGQELNPESYAIKADMLIKGPQVDNIVFGNTLSDDHAA
KFDYMLSNPFGVEMWKIVRAEAEQQGYNGRFGPLPRVSDGSLFLHLLSKMR
PAVDGSRFGVLGSLFTGGAGSGEIRRHLESOLVEAIVALTPTDMFYNTGAT
YVMIVSNKRGKLVQGLQIDAGFWRMRKSLGSKRKEMGEDDIAAITRLFGAFVEA
ELASVDAEGKVDKIVVAGSAPPAPEGVKVLAPLSKILPNSARGYRTITVERPL
VDEAGKVLGKGNKCKPOADSAARDTENVPVSEVETVFAREVLPHPADWIDADK
TKGYEIPFNRHFVVFEPDRDQADLRAVTDQIKAMIAELAA"
2056. .3408
/gene="CC0621"
2056. .3408
/gene="CC0621"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=1
/product="type I restriction-modification system, S subunit"
/protein_id="AAK22606.1"
/db_xref="GI:13421830"
/translation="MSPPAYESKESGVENLGRVPSHWNFRPLKULVIMRSGGTPSKE
REDYWGGEIPWASAKDLKVDTLTDQHLTAELDEGAQQLPANAVVILVRGKMLAR
TFPVLRSRPTINQDLGLIANRGVDPNLYLSRASEVETLCRLDEAGHGTKALRM
DAWSTWELPAPLAEQQAIAAFIDRETAKIDALVEAQERLIALLKEKRAVISHAVTK
GLDPSAQMDKSGVLEQMPAHVEVVPKKNLADSIKAGPFGSALTMDMYSSAGRVYVG
QEOVIGDFRIGDYVYTSDRYNELSOYRVVEGDLLVSCVCTGCKIAIFPQAGBGLIN
PRLIRPNQNPQPTLYLCVLLRSVAFESQSYLSRGTMVDVINGILGEIVVPPPNQ
EQISAGYLAEOEQFDLSAASEAAITLLQERRAALISAAVTGKIDVRLGVQTAEA"
3405. .4448
/gene="CC0622"
3405. .4448
/gene="CC0622"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAK22607.1"
/db_xref="GI:13421831"
/translation="MTRLMIIVRAGRIGERELEAITQGRLLPGFGEVADLTGLKDREAI
LRVLVAITPEGATURNFAAQMNFVNTIGVGLDVVLPRKVTSGVAIGRVTPGVRHD
PAAPIAHTRAVEWLEAVPREAFAQDLRHSFGAFPLTCEVKNREALARVAELTKGD
PGPLGKGVTVKPAVEDAEEDVELDVEDIARQOIIALIKSAFAGHDLAYLVAEIL
RIEYVTVKSPGPGGVDTILAAAGRLGLEDICRIVQVKGDAPANQEVVILKMGVQ
AAKARTGLVLSIAGVNGPAQLIDSEFFSVRLWQPELLTALTALFASQGSDETRAKIP
LKQIWPVAGVDA"

gene
CDS
gene
CDS
gene
CDS
gene
CDS
gene
CDS

4459. .7353
/gene="CC0623"
4459. .7353
/gene="CC0623"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAK22608.1"
/db_xref="GI:13421832"
/translation="MSLHKEIRFEDEICAHLAANGWLVDGAAGDRHDRTRALYVEDLV
AWYQDSQPEAMDGLSRANGAAQVLADRVKSLDERGLDVLNGLVGLVARGRIOM
AQPKPMNPPEIVRRYEANRLRVARQVKNVGENCLDLVLYNLGLVPVATLAEKSD
TQSVEDAINOYRPREPKGREEPPLAFPGGALVHFVAVSNSEVAMTTKLGRODTFF
LPNKGDHGKGNPLNPNHRTAYLWEEIWARDSKLEILGRMYMTKRAKXRTISVWF
PRYHOLDATKLAQVLAEGAGKFLIHSAGSGKTSIAWSAILSELDHADKDKLF
SSVIVSDNRVDTQLOEALDFORTAGVATIKISGSGSGOLAALAAAGKKTIVCT
IDTFFPALEAVREQAQTQKAFVIADEAHSSQTGEAGKLQMLSDAEIAELADGGE
IQTEFTQNTANMAGRAADKSTIYVAFATPKAKTLQLPGRPDAGLPAPFHVYSRQA
IEEGFILDVKNVTPYSLAFKLANDGOEWDEREVERTALKGIMRWVRLHPVNIQKV
QIVVEHYRENVQPLLDCKAKAMVVLGSRQEAVRWSVAINVIRKGYPLGTLVAFSGE
VNDELGAVSETSAALNPGNLNGRDIRDAFKAENPHFLLVANKFTGFPQLCCGYVD
RRLAGIQAVOTLSRLNRAHPGKOTTYILDPCNSAEDILTAFRTYETLAELEAATDNL
IYDLRAKLAAGHYDDHEIERVVRVELDKAKQGDILAALEPVAADLVKAFAAAKQAW
VTKAODDERAAEALDHMNLDLFKGDMAAFQRLYSFLSQIFDYGNTIEKRSIFFR
RLPLLEFGRERCVDLSGLRLTHHALKTKGQDOLLYGTGEAPKLSGMAEVSGLLHE
KEXARLDEIITAKLNDLPEGEITEENCITSSESVLRQPGCGFSGGLASSGFDI"
complement(7350. .8330)
/gene="CC0624"
complement(7350. .8330)
/gene="CC0624"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=1
/product="ISCC3, transposase OrfB"
/protein_id="AAK22609.1"
/db_xref="GI:13421833"
/translation="MRYPACEKIEIIRLQSTLPLVRRTLAQIGVPBRTFFVRWVDLYO
TGGPEALDEDRPAPSRVVMNRIPEDEVRAILLEMALEHSELSELAVRFTDSQGFVSE
ASVYRLKSHDLITSPAFIVMKAATKTKTAPNQMWDITFTFKIIGWGWVLTSTI
LDDFSRYVIAWKLTCTMKAGDVTDLTALSASGCDQSVVRHPRLLSDNGSSYISGE
LADWLKGQEMKHVRGYPHYQTQTKIERHQTILKNRILLENYFLPDLQEAQIGAFIEH
YNHARYHSELENTVPADAYFGRADILRRRASVKRQTLHRRRLQHRKIAA"
complement(8327. .8704)
/gene="CC0625"
complement(8327. .8704)
/gene="CC0625"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=1
/product="ISCC3, transposase OrfA"
/protein_id="AAK22610.1"
/db_xref="GI:13421834"
/translation="MRPKPLAAKATPGERVVKDIRRTRRHFSAECKIRIILDGLRGE
DSTAEILCRKEGIAQSLYVYVWSKFEFMEAGKRRLAGDTARAATTDVEKDLRRESTALKEV
VARQALEIRLLKKSMIADGGGE"
complement(8745. .9119)
/gene="CC0626"
complement(8745. .9119)
/gene="CC0626"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAK22611.1"
/db_xref="GI:13421835"
/translation="MADRKPGDSVMIMNETPPTIVFIGDHPDPSLWDLPWINRAEVAP
AGGRDRLVKLFRGGENNELVSSALPFLPDRPHDLVAINYGEGAPQVFWHTLRFV
EAYRIDFDCPPSAPMQIEGVS"
complement(9136. .9471)
/gene="CC0627"
complement(9136. .9471)
/gene="CC0627"

```
/note="identified by Glimmer2; putative"
```

NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DDBJ accession no. and RGP clone ID.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level, such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.

The orientation of the sequence is from T7 to SP6 of the PAC clone. This sequence of P0656G04 clone has an overlap with QJ1174.D05 (DDBJ: AP003118) clone at the position 1 to 70,328 of 5' end. The sequence of this clone starts at the position 58,198 of QJ1174.D05. This sequence of P0666G04 clone has an overlap with P0489A01 (DDBJ: AP002484) clone at the position 119,689 to 141,983 of 3' end. The sequence of this clone ends at the position 19,295 of P0489A01. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://wq.dna.attfc.go.jp/GenomesGp.html>.

FEATURES
SOURCE

```

source
1..141983
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0666G04"
complement(10870..12330)
/gene="P0666G04.1"
complement(10870..12330)
/gene="P0666G04.1"
/codon_start=1
/product="putative glycin-rich protein"
/protein_id="BAB32899.1"
/db_xref="GI:13122418"
/translaton="MANRREVLVLLMAAVVAVVYVDHRDPAVAVWVRQSMVVGEGG
CGDKHSTATISGGVGVRGGGCVGGVGGAGTGGCAAGLGGGGGGGGGGGGGGGG
LGGGGMGSGSGFGGGGGGCVGGAGFSGSGGVAGGGLRGGGGVCAGGGGGFGGGA
LGGGIGAGSGFGGGAGAGCGVGGGRFGGCGMGGGGFCGGAGGCVGGGGLGCGGM
GGSGPFGAGGAGTGAGGCVGGGIGVGGMGGAAGGIGGAGGCGMGDITGGAGAGGV
GGGGGCMGGGCGFGGGGCGVGGGNTGRDFDGGKNGNLNGDGA SGGITGGASGFGGGA
GIGEVNGFVGAGATGIGLGGSGFGMSGAGGFGTGGSGFGGGLCGGGIGGGGLV
GGGTGCGSARBNAGMCGGTSAGAGGGLGVGGGGAGGCGIGASRCARGCSGAGGCA
GMGGVCGARGGGEHKRGKGKHGGSGCVLGRFI"
complement(join(13539..14009,14482..14609,15099..15240,
15289..15675,15815..16000,16152..16429,16527..16653,
17016..17212,17380..17427,17504..17537,17607..17735,
18635..18730,19189..19255,19524..19618,20292..20387,
20897..21992,22991..23172))
/gene="P0666G04.2"
complement(join(13539..14009,14482..14609,15099..15240,
15289..15675,15815..16000,16152..16429,16527..16653,
17016..17212,17380..17427,17504..17537,17607..17735,
18635..18730,19189..19255,19524..19618,20292..20387,
20897..21992,22991..23172))
/gene="P0666G04.2"
CDS

```

```

/cocon_start=1
/product="putative DNA mismatch repair gene"
/dbn_id="BAB32900.1"
/dn_xref="GI.13122419"
/translation="MQPRRRQQQQQSIILSLQKKPAAAGEGATPRPPRPPAASVA
GIWMLVRPQQQQQDQDASQARHVEQKQVRLKVNQKLTSTNECSALFLFESCKGNK
AAMLIAEGSNMTPLOELKSLWPSKDQFVRASTLSPELGLQNLPHCPCKLPFESS
NNNICIGANSFEEDVQTPDQTSKTPVWESSRGADTPLESDSDQTPLOHPSKFVS
PNEGVRGATLFLDALDSNTPRRSESSKLSGSPSDLPYIKATKLTFTFEDSGVTSQNH
KLSSVKNPKDRICAGALESEDFDPSPLKPTFVMRAVIPLKRVQEDQDRVLTNDSCS
PFWPQNKVKPAQCSVPENKVKHMSAEARSCEWLNPSNIIDANRRRLADPLYDKTT
LFTPPDALRMKSTSQKYNNIKCKYMDWVLFFVKYKFELYEVDAGEIGOKELDMKWITI
SGVQKCRQIIEVNMGLLITVPTRDANNLNLHQDLNLDGKVRGLEQMESADQAKSR
GNSVLRLKLVHVTSTPTSDGNSIGADAVHLLSKETILASNGSRVYGFALDYAAAL
IWTGSHDDQTPFAALGALLVQSPKEIYIETSLGSKETHRLIKKYASAGASCKMOLFTPL

```

```

/note="identified by Glimmer2; putative"
/codon_start=1
/transal_table=11
/product="hypothetical protein"
/protein_id="AAK22612.1"
/db_xref="GI:13421836"
/translation="MGLGMSANNAAPSARYPTSKWPVADLIILRVCAADLDRWVTVRCGP
HCGIAROLFPELAAKRLADVPYIKLIFERGAFCRKAQKCGNGVPASEISVDAMDVGQ
LQNVACMSR"
gene      complement (9893. .9969)
          /gene="trNA-Arg-4"
tRNA      complement (9893. .9969)
          /gene="trNA-Arg-4"
          /product="trNA-Arg"
BASE COUNT      1874 a 3295 c 3142 g 1726 t
ORIGIN
Query Match      48.5%; Score 25.2; DB 1; Length 10037;
Best Local Similarity 90.0%; Pred. No. 1.7e+03;
Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy      23 TCGCCTTCGTCGTCGCCGCCCTCGCCAGTG 52
Db      1311 TCGCTCTCGCCGCTGCCGCCGCCGCCGCGTG 1282

```

RESULT 34	AP003047/c	LOCUS	DEFINITION	AP003047	141983 bp	DNA	linear	PLN 29-MAR-2003
ACCESSION	AP003047	ORyza sativa (japonica cultivar-group)	genomic DNA, chromosome 1,					
VERSION	AP003047	PAC clone: P0668G04.						
KEYWORDS	AP003047.2	GI:13122417						
SOURCE	ORyza sativa (japonica cultivar-group)							
ORGANISM	ORyza sativa (japonica cultivar-group)							
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Orzyzaceae; Orzyza.							

1 *Animalia*: *Arthropoda*, *Oryzidae*, *Oryza*.

REFERENCE
AUTHORS

Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., Katayose, Y., Wu, J., Niumura, Y., Cheng, Z., Nagamura, Y., Antonio, B.A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K., Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K., Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y., Idonuma, A., Iijima, M., Ikeda, M., Ikeno, M., Itoh, S., Itoh, Y., Itoh, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S., Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maebara, T., Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M., Nakama, Y., Nakamichi, S., Nakamura, M., Namiki, N., Negishi, M., Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T., Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Teuji, K., Waki, K., Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K., Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J. H., Kim, H. I., Eun, M. Y., Yano, M., Jiang, J. and Gojobori, T.

TITLE
The genome sequence and structure of rice chromosome 1
Nature 420 (6913), 312-316 (2002)

REFERENCE AUTHOR'S TITLE JOURNAL	COMMENT
2 (bases 1 to 141983) Sasaki, T., Matsumoto, K. and Yamamoto, K. Direct Submission Submitted (13-DEC-2000) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kamonndai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/ , Tel:81-298-38-7441, Fax:81-298-38-7468)	On Feb 23, 2001 this sequence version replaced gi:11862977. Genes were predicted from the integrated results of the following: GENSCAN.0, BLASTX2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI NonRedundant Protein database, nr (http://ncbi.nlm.nih.gov/blast/db/) and the cDNA sequence database at

```

YGLYFSDVSEIOTLIDSRGVFKASTSSWLSALNSSVNKDVAICALGGLVSHLTRLMLE
DALKNCEVLAYHYVRTCLRMDSQTLVNLIEFSNNFDGSSGTYLKHNLNITPCGKRL
LRWVICHPLKDIDAINERLDIVVEGFIQNGCLGVSVTLEHLRKVPDLERLGLKVKSTVGL
SSAVLLPFVGEKILKRIKTFGLMVKGLRGVLDLLILQRODHGALSANVDIPTLS
SUGELTHFEAEIADDDFPYQDHSVKDDDDANTLMLVLLVGLKASWSLVINALSTID
VRSFAMALSSFGTTCRNNILKKGAPVLQMKGLWHPYAFASVNGVLVNDLSLQON
LQSNRFAILLTGPNNNGKSTIMRATCLAIIVLAQSLFECIIIEVASDGLGCVVPCQS
CELTADAIATFRIGAMDRIMSGESTFELVCTETASILENATEDSVLLDELGRGTSTF
DGVAIAYVFRHLVAVRCLLPFATHVPLTKEFASHPHVTLOHMACMLKPRNGDGG
EKELFTYRLTSGCPESYGLOVATWAGLPSTVERASAGEMMRSKIAGNFRSSER
AFESTLHEWRTIVAGVKAHDLDETHDTLFCVFHELKAHFRKR"
complement(join(24484..25102,25464..25506,25942..26053,
26323..26331,26395..26557,26679..26785,26919..27000,
27084..27121,27204..27290,28128..28238,28380..28517))
/genes="P0666G04.3"
complement(join(24484..25102,25464..25506,25942..26053,
26323..26331,26395..26557,26679..26785,26919..27000,
27084..27121,27204..27290,28128..28238,28380..28517))
/genes="P0666G04.3"
/notes="hypothetical protein
similar to Arabidopsis thaliana chromosome 1 F14D16.10"
/codon_start=1
/protein_id="BAB32901.1"
/db_xref="GI:13122420"
/translation="MASPAKRARPASSVPDPASPAPEDPCVLLRRSWELASVLHF
LRVFEVIEGNLGLSAEIEETALIANHDLARIHIALIKGIPPTIKNLKVDGMITAT
AKKLSQDMTWAEAGNPFKIPGKEIDTVKQDPIKRLILKALCEVRSESDAVWVY
NDMKRGVDISNPRKEKLGSSNGTIYVYDGSTIGHRLYSEVTLDPFRNLKGNGR
LTKPVNIQETVATNLDSEIASEKLSKRPESAIGELHKTETIIPDEKLOKKGR
DUKROQKDELLAYATSFTQRLRRRRPYSYNSDYRSIKEAKAAAIHVLECRSK
ESDQLSSDEDEDRYNGKDDSGSDGSDGENDAYDPNKGDTTEEEVFVPRK
RTRLASRRFDKPOGLRRRRRNKNDENVRPQLTPRSMTKTKMRQRTSISKQFSL
SGSEDDREMIADSEESGSP"
join(33254)..34423,35296..35505,35612..35707,35804..36025,
36599..36793,36867..37013,37093..37242,37532..37662,
37748..37964)
/genes="P0666G04.4"
join(33254)..34423,35296..35505,35612..35707,35804..36025,
36599..36793,36867..37013,37093..37242,37532..37662,
37748..37964)
/genes="P0666G04.4"
/notes="contains ESTs
D22091(C10325),C26080(C11596),C26176(C11773),
AU085955(EC2104)"
/codon_start=1
/product="putative heat-shock protein"
/protein_id="BAB32902.1"
/db_xref="GI:13122421"
/translation="KSVVGFDLGNESCIVAVARQRIQIDVVLNEESKRETPAIVCFDGD
QRFIGTAGAASATMPKNSVQIKRLGLRKYSDPELQRIIAAPFRFVSGPDPFLVH
ARYLGEARVFTPTQLMAMVLSNLKGLAESNLNTAVVDCIGIPVFTDLQRAVLDA
TIAGLCPLRFHETTATLAYGIYKTDLPENQUNVAFVDVGHASMQVCVYKKGQAL
KLISHAYRSLGGRDFEVLFFHFAKFKDEYKIDVYQARACILRVACEKLKKVLS
ANPEPMHIECLMDKVRGFTKREEFEKISILERVKGLPEKALAEAGLTENVHF
VEVSGSRVPAIKIITLDFGKEPRPTWNSCEVARGCALCEALISPTFKVREFOVN
DGPFPSIAMSQVPSQNGNQTVFPKGNPLPSVKALTFYRSNTFOVDVTVDTGDL
QISPKISTVTGVPNPGKDKAKLVKVLNLIHGVTVESATMLEEEVEVPAATTE
PKDSAKMETDAPNEASGTDVNWQEKAKADAAAGNAPNMESEKSVPMETDAK
VPSKRRKKTNPVAELVYGLTTELOKAVEYEMALQDRVMEETKDKGNVSEY
VYDMRNLKYDNKVPFDVDEKFAFTAKLOVEDWLYEDGEDETGKYVYAKLEELKVG
GPTEARYEWRMGDSIDQLAYCINSFDRALSKDPKFDHIEMEKQKVINCSAEV
WLREKIQOQDALPKHANPVLSSDLKKKAEVTDRPCKPTMMKPKAPKQPTPTPPT
ETPAGGAOTPEQOPQGAAGAEGSGASESTGEOMETDKPEGTEA"
join(42951..43285,43715..43762,44108..44284,44440..44596,
44728..44811,45332..45387)
/genes="P0666G04.5"
join(42951..43385,43715..43762,44108..44284,44440..44596,
44728..44811,45332..45387)
/genes="P0666G04.5"
/notes="hypothetical protein
similar to Arabidopsis thaliana chromosome 3 MRC8.21"
/codon_start=1

```

```

/protein_id="BAB32903.1"
/db_xref="GI:13122422"
/translation="MPLPRAATVIRRGQIHRVTAMPPPQQAPEPAGLLOLLOPEV
VAPNPPPPAPPAPAPPPQGEARYRRLVRLQAVPNEDHVPDNYGDPDELGIT
PAYQALEHLPDLAGAPAEVKRYEMRSVLRNYPSPQRIQRTONREYRERILSAY
QPHLELYNDPSTFTILPAFLAINGNTTESITSIMMEPAGVFAFPMLKPSFCQMLM
SEVNNFLRWAQANORIMRPTSLDRHGRGAALSDFLQELMDLNMKDFISPMSTVLPF
EYVGNTLDSHHTFVLEYGEADGARAIQNCAAHTIEDILVFCC"
complement(join(46556..46681,51622..51919,52095..52141,
52141..52141))
Query Match 48.5%; Score 25.2; DB 8; Length 141983;
Best Local Similarity 71.7%; Pred. No. 1.4e+03;
Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 7 TTCTCTCACGCTCTCATCGTCTCTCGTCGCTGCGCGGCTCCCGCAGTG 52
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97528 TTCTCTCACACTCCGATCGCTCTCGCCATGGACGCGCTCGGCGCG 97483

RESULT 35
TBCHRIA2
LOCUS
DEFINITION
Trypanosoma brucei DNA chromosome 1, segment 2/4.
ACCESSION
AL929604.1
VERSION
AL929604.1
KEYWORDS
GI:25991392
SOURCE
Trypanosoma brucei
ORGANISM
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE
1
AUTHORS
Hall,N., Berriman,M., Lennard,N.J., Harris,B.R., Gerrard,C.S.,
Ackn,R.J., Barron,A.J., Bart-Delabesse,E.N., Bowman,S.,
Bray-Allen,S.P., Bringaud,F., Clark,L.N., Corton,C.H., Cronin,A.,
Davies,R., Doggett,J., Fraser,A., Gruter,E., Hall,S., Harper,D.A.,
Hertz-Fowler,C., Kay,M.P., Leech,V., Mayes,R., Price,C.,
Quail,M.A., Rabinowitsch,E., Rutherford,K., Sasse,J., Sharp,S.,
Shownkeen,R., Gull,K., Barrell,B.G. and Melville,S.E.
The sequence and analysis of the highly polymorphic chromosome 1 of
the African trypanosome, Trypanosoma brucei
Unpublished
2 (bases 1 to 281000)
Berriman,M., Hertz-Fowler,C.V.A., Hall,N., Kerhornou,A.X.,
Bowman,S., Quail,M., Kay,M.P., Bray-Allen,S., Lennard,N.J.,
Clark,L.N., Harris,B.R., Melville,S., Gerrard,C., Rajandream,M.A.
and Barrell,B.G.
Direct Submission
Submitted (20-SEP-2002) The Wellcome Trust Sanger Institute,
Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
For more information about this sequence or the Trypanosoma brucei
Project, see http://www.sanger.ac.uk/Projects/Tbrucei. NOTE: This
is the largest contiguous sequence from chromosome 1,
spanning between the 70 bp repeat arrays at either end. The
total
size is 1056003 bp but it has been split into 4 database
entries
TBCHRIA1 - 1..197000
TBCHRIA2 - 197001..478000 (THIS ENTRY)
TBCHRIA3 - 478001..807000
TBCHRIA4 - 807001..1056003
It replaces the previously submitted unfinished (941631 bp)
sequence for T. brucei chromosome 1.
FEATURES
Location/Qualifiers
1..281000
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/strain="TREU927"
/db_xref="taxon:5691"
/chromosomes="1"
complement(1000..1821)
/gene="Tb927.1.530"
complement(1000..1821)
/gene="Tb927.1.530"

```

variation	<pre>/codon_start=1 /product="hypothetical protein" /protein_id="CAB95347.1" /db_xref="GI:9366585" /translations="MRSRYMRKGNESVTVAVSFTRDAFFLSCMRSLGIGVAPYLRRC RSPQDFPLKFLSTLAETPENRI VAAVPCTTVMTVDDHDDVDGMLPPLQSCQAC SPSHVSHFOLLVLSLFAIQACRTTSTSSKSHQROLSPATFKSDVEDAAATFLQI LSNSVPPNELMNCRYQTQTHSCRLTKDRMLKNLKEGPVLAUPLVDLMISQNVAE GNVALRKCDARQSLMRSLNPLKCAKOHLSVCDDDAAAYLULETTVAVKYEIPLNL" 2218 /notes="Majority of WCS, TPB15A20, TP7E4 and TPB12C6; reads call T, Two TP20A12 reads call C, type 1 (single nucleotide substitutions)" complement(2222..5611) /gene="Tb927.1.540" complement(2222..5611) /gene="Tb927.1.540" /EC number="2.7.7.6" /codon_start=1 /product="DNA-directed RNA polymerase III, probable" /protein_id="CAB95348.1" /db_xref="GI:9366586" /translations="MTPRARGCDGYATRKALLPALINMKGLNHHVASFDDLINVE LGRILNESNREIKSSIDPDFISIRYGI RVCRRPREIVGKQLPKPSPHECRTRDLTY RGDI VDVVEYTSRDSKRICTETNI KIGTIPIMLKSSCNLYGKGREELTAMRECPID PGCYFI IKGVEKVLQEOQSKNRVII EADDNGDI VAHVQSKTHYSISKCVTRFKGR IVMHRSPKEDIPIVIVVLKALHESDQIIAQHIGVSPAFOSVLPACFEHAASLNVKIQ DDALHFIGERRKETFWEMDETQQRHVOKSKADSAEFLANVLGLCHIREGEVQRDMNFR HKALYVCFMVRVEASVSSLDLDERDFYGNKRFEITFTGLMLGLLFEDLLKOYRNVVKT AMQLLASIKDHTKFPNLKQMESKMEVIQNGMRVLSASSGKRWELKRFNMSRQGITQVLS RLSVISICIGMTPLASFEKSRIKISGRSLQPSQMGWVCPCDTPEGESCLGVKNFAL SOYTLMDSDSFVRAAAVNLGVEVDICI SPNEFLKYTVFLNGTLMGHRYPSRLRGV RTLRRSLGPHVSIALNDRQKSVQICDGRVRLYIVVRDKPVTISOHLNLSG KCSINFLAGLVEFDINEANDCLIAVYNOYIRMDTVMLGAYPQRLCTRKAMNLNYEIL HHNOSARNTFOSAMGQALGTIAYNOYIRMDTVMLGAYPQRLCTRKAMNLNYEIL GAGINAMCVMSYSGYDIEDAQVYNNRSLDRGYRCVLRKHEVELERLPNGSYDIIL PADSNVGRSHALSADGIASKGAVVRQFDILVNKYTPVPSMEPRNPVLVYKYPQAVVD HVVISPPGSDSRMSDEOKI KVTTRVREPGPDGKFSRHGQKGVGLITIDGYNMFPN EOGICPDMINMHPFSPMTVGLLELVSSKVALTKGKFGDTAFGCDKAEDLGEELI RHGYNVHGKDI FVSGVTGEMLRAYVFGPIYIYORLKHMTDKMHARATGRPSMLTROP TEGRSRGUGMERKDCMWGYSASNLNLERLLVSDLFTIDICRCCGPMGYNGVCS 3257 /notes="Majority of WCS, TP3D18, TP7E4 and TPB15A20; reads call T, Two TP20A12 reads call C, One WCS read calls C type 1 (single nucleotide substitutions)" 3343 /notes="Majority of WCS, TPB15A20, TPB12C6, TP3D18; and TP7E4 reads call C, One WCS read calls T type 1 (single nucleotide substitutions)" 3734 /notes="Majority of WCS, TPB15A20, TP3D18 and TPB12C6; reads call A, Three TP20A12 reads call G type 1 (single nucleotide substitutions)" 5345 /notes="Majority of WCS, TPB15A20, TPB12C6; and TP3D18 reads call A, One WCS read calls C type 1 (single nucleotide substitutions)" complement(5981..6283) /gene="Tb927.1.550" complement(5981..6283) /gene="Tb927.1.550" /notes="unlikely gene" /codon_start=1 /product="hypothetical protein, unlikely" /protein_id="CAB95349.1" /db_xref="GI:9366587" /translations="MDSFCALCLTFCIDERGSRAPCLHSIIWVKHFWFHEGWADDEE TPRFELHIFRDATAFVQMAVMGDFNLLYVGRASSGWSVGKKEKKTETLMLHYVS" complement(6507..7835) /gene="Tb927.1.560" complement(6507..7835) /gene="Tb927.1.560"</pre>
variation	<pre>/codon_start=1 /product="conserved hypothetical protein" /protein_id="CAB95350.1" /db_xref="GI:9366588" /translations="MLFVASSSSKSHLINLDTKTVERSYTPTVERRGLVYMPARTAVVM SNORSCASIFSPAMOQPVLCITPOAATASASTSCGAFLLVGTAGCNLYVWSTLSGE LLHVRSHRRITTEIALISDQSLIATASESVCKTWLLAALVARGVSPAPRIVPNGH LSLVNACSPLDSSSVTASSDRSKRIFDGTGQQQVVTLDGVLTSVRSPSGDEMLL IGGETSLFVRVLYSFRSQCLTAGLIRCVNDVVTCTPFTDGHKGAIVFIHFDYTRP DYVLVSGTGVILMNIIRTAATVEKAVDDIAGGLLSVCYVPSDATOPSSAAVPPVILQ KHPLDPLGVDFIVVSGAKESAPISYKEGGSRRKRRKHQOQTMEDTCVSMKRTENV VATELEGARESTGKPTSAERFSRLREKQDELEFLKNLKEKLRKLAAS" 7583 /notes="Majority of genomic shotgun, TP7E4, TP3D18, TPB12C6 and TPB15A20 reads call A, One genomic shotgun read calls G type 1 (single nucleotide substitutions)" 7896 /notes="Majority of WCS, TP3D18, TP7E4, TP3D18 and TPB12C6 reads call G, One WCS read calls A type 1 (single nucleotide substitutions)" complement(8396..9223) /gene="Tb927.1.570" complement(8396..9223) /gene="Tb927.1.570" /codon_start=1 /product="conserved hypothetical protein" /protein_id="CAB95351.1" /db_xref="GI:9366589" /translations="MDARVHRSKRGCGRVLKRSSPSEVVAEVGTCRPGDPFVLPRLR SPNGSGCHHCPCDCLYVSGLELCRYLSQVRSQVTCGRRELCAKSVLPFPFVAR LLRVANVADDTFTFDUGGVSVLFHVALATGASCVCVEINEHNAKVAKEAHTLRPV FRRKRKGLDVSIVCGDFCKVODNYFSSSCVVMANLLMPRVFNHLYSERLSRPI GSRLVCMEDLYPHRSVAAARDPAFEKFMVDYRWQEDSVSEVSPASGPFYLYKRS" complement(10259..11809) /gene="Tb927.1.580" complement(10259..11809) /gene="Tb927.1.580" /codon_start=1 /product="phosphate-repressible phosphate permease, possible" /protein_id="CAB95352.1" /db_xref="GI:9366590" /translations="MWHPLYLWIIIVGIVSFLTCGCGVMNDLANSFGTTVGSRVNLNW QIVILASICEFVGAVSLGSEVTSTGGIANPMTFAPEVPLYVMGMCAISATFIWLL FATMPSLVSSTHSIAGIIFALVGGFGAFKIDFPPVYTVGAPIASWFI5P MFAGAAVAAIYALLRLVLRPANSVNRALPALPIVGTFTFLESFVFLKGAARLHW GPAKASVAAALIGLGAASISAACIPIILRRRVRLITERAERAEETGMNTAPEISGDAG AISENAAGVGAAVEGPDVTANRIVPPSSEPTSDPTAKENPANNASGLTTPGVVDEAL KEDVOIYDERVEYVRYLQVFTAACASFAGANDVSNAIAPFSMAYSIYINQOVVEEN DVPILVILVAGLVGLVGLIMRLGERITKTPTRSGFSAELSVLVVLSUCSAP GIPVSTHCTITGAVVAISIMDCGFRKVRMMVMGKVLGWIFTLLITAAISALLPAQGI YAPSLT5Q" 12152 /notes="Majority of WCS, TP3D18 and TP7E4; reads call C, Two WCS reads call T, type 1 (single nucleotide substitutions)" 13074 /notes="Majority of genomic shotgun, TPB15A20, TP3D18; and TP7E4 reads call A, Two TP20A12 reads call C type 1 (single nucleotide substitutions)" complement(13430..13846) /gene="Tb927.1.590" complement(13430..13846) /gene="Tb927.1.590" /codon_start=1 /product="hypothetical protein" /protein_id="CAB53050.1" /db_xref="GI:25991393" /translations="MNAFGLCHDHWKRLSRVLLFPFIFSRQLIILALSVPYGTVLI FPPSPFFLFFPFFFLCFIPLSCLSFNLSVLNHTRTAFAPSAAMIPTLWHSVARR VALSGGVEGCRVTDVRECLKQNLIMLEKVGSSSSSR" 13700..13709</pre>
variation	<pre>/codon_start=1 /product="hypothetical protein" /protein_id="CAB95347.1" /db_xref="GI:9366585" /translations="MRSRYMRKGNESVTVAVSFTRDAFFLSCMRSLGIGVAPYLRRC RSPQDFPLKFLSTLAETPENRI VAAVPCTTVMTVDDHDDVDGMLPPLQSCQAC SPSHVSHFOLLVLSLFAIQACRTTSTSSKSHQROLSPATFKSDVEDAAATFLQI LSNSVPPNELMNCRYQTQTHSCRLTKDRMLKNLKEGPVLAUPLVDLMISQNVAE GNVALRKCDARQSLMRSLNPLKCAKOHLSVCDDDAAAYLULETTVAVKYEIPLNL" 2218 /notes="Majority of WCS, TPB15A20, TP7E4 and TPB12C6; reads call T, Two TP20A12 reads call C, type 1 (single nucleotide substitutions)" complement(2222..5611) /gene="Tb927.1.540" complement(2222..5611) /gene="Tb927.1.540" /EC number="2.7.7.6" /codon_start=1 /product="DNA-directed RNA polymerase III, probable" /protein_id="CAB95348.1" /db_xref="GI:9366586" /translations="MTPRARGCDGYATRKALLPALINMKGLNHHVASFDDLINVE LGRILNESNREIKSSIDPDFISIRYGI RVCRRPREIVGKQLPKPSPHECRTRDLTY RGDI VDVVEYTSRDSKRICTETNI KIGTIPIMLKSSCNLYGKGREELTAMRECPID PGCYFI IKGVEKVLQEOQSKNRVII EADDNGDI VAHVQSKTHYSISKCVTRFKGR IVMHRSPKEDIPIVIVVLKALHESDQIIAQHIGVSPAFOSVLPACFEHAASLNVKIQ DDALHFIGERRKETFWEMDETQQRHVOKSKADSAEFLANVLGLCHIREGEVQRDMNFR HKALYVCFMVRVEASVSSLDLDERDFYGNKRFEITFTGLMLGLLFEDLLKOYRNVVKT AMQLLASIKDHTKFPNLKQMESKMEVIQNGMRVLSASSGKRWELKRFNMSRQGITQVLS RLSVISICIGMTPLASFEKSRIKISGRSLQPSQMGWVCPCDTPEGESCLGVKNFAL SOYTLMDSDSFVRAAAVNLGVEVDICI SPNEFLKYTVFLNGTLMGHRYPSRLRGV RTLRRSLGPHVSIALNDRQKSVQICDGRVRLYIVVRDKPVTISOHLNLSG KCSINFLAGLVEFDINEANDCLIAVYNOYIRMDTVMLGAYPQRLCTRKAMNLNYEIL HHNOSARNTFOSAMGQALGTIAYNOYIRMDTVMLGAYPQRLCTRKAMNLNYEIL GAGINAMCVMSYSGYDIEDAQVYNNRSLDRGYRCVLRKHEVELERLPNGSYDIIL PADSNVGRSHALSADGIASKGAVVRQFDILVNKYTPVPSMEPRNPVLVYKYPQAVVD HVVISPPGSDSRMSDEOKI KVTTRVREPGPDGKFSRHGQKGVGLITIDGYNMFPN EOGICPDMINMHPFSPMTVGLLELVSSKVALTKGKFGDTAFGCDKAEDLGEELI RHGYNVHGKDI FVSGVTGEMLRAYVFGPIYIYORLKHMTDKMHARATGRPSMLTROP TEGRSRGUGMERKDCMWGYSASNLNLERLLVSDLFTIDICRCCGPMGYNGVCS 3257 /notes="Majority of WCS, TP3D18, TP7E4 and TPB15A20; reads call T, Two TP20A12 reads call C, One WCS read calls C type 1 (single nucleotide substitutions)" 3343 /notes="Majority of WCS, TPB15A20, TPB12C6, TP3D18; and TP7E4 reads call C, One WCS read calls T type 1 (single nucleotide substitutions)" 3734 /notes="Majority of WCS, TPB15A20, TP3D18 and TPB12C6; reads call A, Three TP20A12 reads call G type 1 (single nucleotide substitutions)" 5345 /notes="Majority of WCS, TPB15A20, TPB12C6; and TP3D18 reads call A, One WCS read calls C type 1 (single nucleotide substitutions)" complement(5981..6283) /gene="Tb927.1.550" complement(5981..6283) /gene="Tb927.1.550" /notes="unlikely gene" /codon_start=1 /product="hypothetical protein, unlikely" /protein_id="CAB95349.1" /db_xref="GI:9366587" /translations="MDSFCALCLTFCIDERGSRAPCLHSIIWVKHFWFHEGWADDEE TPRFELHIFRDATAFVQMAVMGDFNLLYVGRASSGWSVGKKEKKTETLMLHYVS" complement(6507..7835) /gene="Tb927.1.560" complement(6507..7835) /gene="Tb927.1.560"</pre>
variation	<pre>/codon_start=1 /product="hypothetical protein" /protein_id="CAB95347.1" /db_xref="GI:9366585" /translations="MRSRYMRKGNESVTVAVSFTRDAFFLSCMRSLGIGVAPYLRRC RSPQDFPLKFLSTLAETPENRI VAAVPCTTVMTVDDHDDVDGMLPPLQSCQAC SPSHVSHFOLLVLSLFAIQACRTTSTSSKSHQROLSPATFKSDVEDAAATFLQI LSNSVPPNELMNCRYQTQTHSCRLTKDRMLKNLKEGPVLAUPLVDLMISQNVAE GNVALRKCDARQSLMRSLNPLKCAKOHLSVCDDDAAAYLULETTVAVKYEIPLNL" 2218 /notes="Majority of WCS, TPB15A20, TP7E4 and TPB12C6; reads call T, Two TP20A12 reads call C, type 1 (single nucleotide substitutions)" complement(2222..5611) /gene="Tb927.1.540" complement(2222..5611) /gene="Tb927.1.540" /EC number="2.7.7.6" /codon_start=1 /product="DNA-directed RNA polymerase III, probable" /protein_id="CAB95348.1" /db_xref="GI:9366586" /translations="MTPRARGCDGYATRKALLPALINMKGLNHHVASFDDLINVE LGRILNESNREIKSSIDPDFISIRYGI RVCRRPREIVGKQLPKPSPHECRTRDLTY RGDI VDVVEYTSRDSKRICTETNI KIGTIPIMLKSSCNLYGKGREELTAMRECPID PGCYFI IKGVEKVLQEOQSKNRVII EADDNGDI VAHVQSKTHYSISKCVTRFKGR IVMHRSPKEDIPIVIVVLKALHESDQIIAQHIGVSPAFOSVLPACFEHAASLNVKIQ DDALHFIGERRKETFWEMDETQQRHVOKSKADSAEFLANVLGLCHIREGEVQRDMNFR HKALYVCFMVRVEASVSSLDLDERDFYGNKRFEITFTGLMLGLLFEDLLKOYRNVVKT AMQLLASIKDHTKFPNLKQMESKMEVIQNGMRVLSASSGKRWELKRFNMSRQGITQVLS RLSVISICIGMTPLASFEKSRIKISGRSLQPSQMGWVCPCDTPEGESCLGVKNFAL SOYTLMDSDSFVRAAAVNLGVEVDICI SPNEFLKYTVFLNGTLMGHRYPSRLRGV RTLRRSLGPHVSIALNDRQKSVQICDGRVRLYIVVRDKPVTISOHLNLSG KCSINFLAGLVEFDINEANDCLIAVYNOYIRMDTVMLGAYPQRLCTRKAMNLNYEIL HHNOSARNTFOSAMGQALGTIAYNOYIRMDTVMLGAYPQRLCTRKAMNLNYEIL GAGINAMCVMSYSGYDIEDAQVYNNRSLDRGYRCVLRKHEVELERLPNGSYDIIL PADSNVGRSHALSADGIASKGAVVRQFDILVNKYTPVPSMEPRNPVLVYKYPQAVVD HVVISPPGSDSRMSDEOKI KVTTRVREPGPDGKFSRHGQKGVGLITIDGYNMFPN EOGICPDMINMHPFSPMTVGLLELVSSKVALTKGKFGDTAFGCDKAEDLGEELI RHGYNVHGKDI FVSGVTGEMLRAYVFGPIYIYORLKHMTDKMHARATGRPSMLTROP TEGRSRGUGMERKDCMWGYSASNLNLERLLVSDLFTIDICRCCGPMGYNGVCS 3257 /notes="Majority of WCS, TP3D18, TP7E4 and TPB15A20; reads call T, Two TP20A12 reads call C, One WCS read calls C type 1 (single nucleotide substitutions)" 3343 /notes="Majority of WCS, TPB15A20, TPB12C6, TP3D18; and TP7E4 reads call C, One WCS read calls T type 1 (single nucleotide substitutions)" 3734 /notes="Majority of WCS, TPB15A20, TP3D18 and TPB12C6; reads call A, Three TP20A12 reads call G type 1 (single nucleotide substitutions)" 5345 /notes="Majority of WCS, TPB15A20, TPB12C6; and TP3D18 reads call A, One WCS read calls C type 1 (single nucleotide substitutions)" complement(5981..6283) /gene="Tb927.1.550" complement(5981..6283) /gene="Tb927.1.550" /notes="unlikely gene" /codon_start=1 /product="hypothetical protein, unlikely" /protein_id="CAB95349.1" /db_xref="GI:9366587" /translations="MDSFCALCLTFCIDERGSRAPCLHSIIWVKHFWFHEGWADDEE TPRFELHIFRDATAFVQMAVMGDFNLLYVGRASSGWSVGKKEKKTETLMLHYVS" complement(6507..7835) /gene="Tb927.1.560" complement(6507..7835) /gene="Tb927.1.560"</pre>
variation	<pre>/codon_start=1 /product="hypothetical protein" /protein_id="CAB95347.1" /db_xref="GI:9366585" /translations="MRSRYMRKGNESVTVAVSFTRDAFFLSCMRSLGIGVAPYLRRC RSPQDFPLKFLSTLAETPENRI VAAVPCTTVMTVDDHDDVDGMLPPLQSCQAC SPSHVSHFOLLVLSLFAIQACRTTSTSSKSHQROLSPATFKSDVEDAAATFLQI LSNSVPPNELMNCRYQTQTHSCRLTKDRMLKNLKEGPVLAUPLVDLMISQNVAE GNVALRKCDARQSLMRSLNPLKCAKOHLSVCDDDAAAYLULETTVAVKYEIPLNL" 2218 /notes="Majority of WCS, TPB15A20, TP7E4 and TPB12C6; reads call T, Two TP20A12 reads call C, type 1 (single nucleotide substitutions)" complement(2222..5611) /gene="Tb927.1.540" complement(2222..5611) /gene="Tb927.1.540" /EC number="2.7.7.6" /codon_start=1 /product="DNA-directed RNA polymerase III, probable" /protein_id="CAB95348.1" /db_xref="GI:9366586" /translations="MTPRARGCDGYATRKALLPALINMKGLNHHVASFDDLINVE LGRILNESNREIKSSIDPDFISIRYGI RVCRRPREIVGKQLPKPSPHECRTRDLTY RGDI VDVVEYTSRDSKRICTETNI KIGTIPIMLKSSCNLYGKGREELTAMRECPID PGCYFI IKGVEKVLQEOQSKNRVII EADDNGDI VAHVQSKTHYSISKCVTRFKGR IVMHRSPKEDIPIVIVVLKALHESDQIIAQHIGVSPAFOSVLPACFEHAASLNVKIQ DDALHFIGERRKETFWEMDETQQRHVOKSKADSAEFLANVLGLCHIREGEVQRDMNFR HKALYVCFMVRVEASVSSLDLDERDFYGNKRFEITFTGLMLGLLFEDLLKOYRNVVKT AMQLLASIKDHTKFPNLKQMESKMEVIQNGMRVLSASSGKRWELKRFNMSRQGITQVLS RLSVISICIGMTPLASFEKSRIKISGRSLQPSQMGWVCPCDTPEGESCLGVKNFAL SOYTLMDSDSFVRAAAVNLGVEVDICI SPNEFLKYTVFLNGTLMGHRYPSRLRGV RTLRRSLGPHVSIALNDRQKSVQICDGRVRLYIVVRDKPVTISOHLNLSG KCSINFLAGLVEFDINEANDCLIAVYNOYIRMDTVMLGAYPQRLCTRKAMNLNYEIL HHNOSARNTFOSAMGQALGTIAYNOYIRMDTVMLGAYPQRLCTRKAMNLNYEIL GAGINAMCVMSYSGYDIEDAQVYNNRSLDRGYRCVLRKHEVELERLPNGSYDIIL PADSNVGRSHALSADGIASKGAVVRQFDILVNKYTPVPSMEPRNPVLVYKYPQAVVD HVVISPPGSDSRMSDEOKI KVTTRVREPGPDGKFSRHGQKGVGLITIDGYNMFPN EOGICPDMINMHPFSPMTVGLLELVSSKVALTKGKFGDTAFGCDKAEDLGEELI RHGYNVHGKDI FVSGVTGEMLRAYVFGPIYIYORLKHMTDKMHARATGRPSMLTROP TEGRSRGUGMERKDCMWGYSASNLNLERLLVSDLFTIDICRCCGPMGYNGVCS 3257 /notes="Majority of WCS, TP3D18, TP7E4 and TPB15A20; reads call T, Two TP20A12 reads call C, One WCS read calls C type 1 (single nucleotide substitutions)" 3343 /notes="Majority of WCS, TPB15A20, TPB12C6, TP3D18; and TP7E4 reads call C, One WCS read calls T type 1 (single nucleotide substitutions)" 3734 /notes="Majority of WCS, TPB15A20, TP3D18 and TPB12C6; reads call A, Three TP20A12 reads call G type 1 (single nucleotide substitutions)" 5345 /notes="Majority of WCS, TPB15A20, TPB12C6; and TP3D18 reads call A, One WCS read calls C type 1 (single nucleotide substitutions)" complement(5981..6283) /gene="Tb927.1.550" complement(5981..6283) /gene="Tb927.1.550" /notes="unlikely gene" /codon_start=1 /product="hypothetical protein, unlikely" /protein_id="CAB95349.1" /db_xref="GI:9366587" /translations="MDSFCALCLTFCIDERGSRAPCLHSIIWVKHFWFHEGWADDEE TPRFELHIFRDATAFVQMAVMGDFNLLYVGRASSGWSVGKKEKKTETLMLHYVS" complement(6507..7835) /gene="Tb927.1.560" complement(6507..7835) /gene="Tb927.1.560"</pre>

```

/note="Majority of WCS,TP3D18,TP7B4,TPB15A20 and;TP16A7
reads call 10 G, One WCS and TP20A12 read call 12G

Query Match      48.5%; Score 25.2; DB 3; Length 281000;
Best Local Similarity 71.7%; Pred. No. 1.3e+03;
Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Oy  4  CAGTTCTCTCAGCTCTCATCGCTCTCGTCTCGCTCGCGCCCTCGCCA 49
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   64495  CATTCTCTCTCGCTCATCACCGTCATCTCGTCTCTCTCGTCA 64540

RESULT 36
PSU96179
LOCUS      723 bp      DNA      linear      BCT 06-NOV-2001
DEFINITION Pseudomonas syringae pv. syringae Hma (hrma) gene, partial cds;
and unknown gene.
ACCESSION U96179
VERSION   U96179.1 GI:2149939
KEYWORDS  .
SOURCE    Pseudomonas syringae pv. syringae
           Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
           Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 723)
           Alfano,J.R., Klm,H.S., Delaney,T.P. and Collmer,A.
           Evidence that the Pseudomonas syringae pv. syringae hrp-linked hrma
           gene encodes an Avr-like protein that acts in an hrp-dependent
           manner within tobacco cells
           Mol. Plant Microbe Interact. 10 (5), 580-588 (1997)
JOURNAL   97348579
MEDLINE   9204563
PUBMED    2 (bases 1 to 723)
AUTHORS   Alfano,J.R. and Collmer,A.
TITLE     Direct Submission
SUBMITTED (02-APR-1997) Plant Pathology, Cornell University, 334
JOURNAL   Plant Science Bldg., Ithaca, NY 14853, USA

FEATURES             source
     1. .723
     /organism="Pseudomonas syringae pv. syringae"
     /mol_type="genomic DNA"
     /strain="61"
     /db_xref="taxon:321"
     /notes="pathovar: syringae"
     356. .646
     /note="ORF1"
     /codon_start=1
     /transl_table=11
     /product="unknown"
     /protein_id="AAB58695.1"
     /db_xref="GI:2149940"
     /translations="MIDKAFALTLLRDDTHQRLLLTGLLEPHEDLPLORLLAGALNP
     LVNAGPGIGHDEQSLGYHAYQSIPREKVSVMELKLEIAGLVEMKWCWEART"
     706. >723
     /gene="hrma"
     706. .>723
     /gene="hrma"
     /notes="Avr-like protein"
     /codon_start=1
     /transl_table=11
     /product="Hrma"
     /protein_id="AAB58696.1"
     /db_xref="GI:2149941"
     /translations="MNPPIHA"
     170 a 202 c 173 t

BASE COUNT 170 a 202 c 173 t
ORIGIN

Query Match      48.1%; Score 25; DB 1; Length 723;
Best Local Similarity 69.4%; Pred. No. 2.4e+03;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy  4  CAGTTCTCTCAGCTCTCATCGCTCTCGTCTCGCTCGCGCCCTCGCAGTG 52
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db   464  CTGTTGGCTGCGCTCTCAACCCCTTGTGAATCGCGCCCGGCATTG 512

RESULT 17
AF232003
LOCUS      2898 bp      DNA      linear      BCT 23-MAY-2000
DEFINITION Pseudomonas syringae pv. syringae strain 61 trna-Leu gene, complete
sequence; putative type III chaperone and type III effector protein
(hrma) genes, complete cds; and hrpK (hrpK) gene, partial cds.
ACCESSION AF232003
VERSION   AF232003.1 GI:8037767
KEYWORDS  .
SOURCE    Pseudomonas syringae pv. syringae
           Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
           Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 2898)
           Alfano,J.R., Klm,H.S., Delaney,T.P. and Collmer,A.
           Evidence that the Pseudomonas syringae pv. syringae hrp-linked hrma
           gene encodes an Avr-like protein that acts in an hrp-dependent
           manner within tobacco cells
           Mol. Plant Microbe Interact. 10 (5), 580-588 (1997)
JOURNAL   97348579
MEDLINE   9204563
PUBMED    2 (bases 1 to 2898)
AUTHORS   Alfano,J.R., Charkowski,A.O., Deng,W.L., Badel,J.L.,
           Petnicki-Ocwieja,T., van Dijk,K. and Collmer,A.
           The Pseudomonas syringae hrp pathogenicity island has a tripartite
           mosaic structure composed of a cluster of type III secretion genes
           bounded by exchangeable effector and conserved effector loci that
           contribute to parasitic fitness and pathogenicity in plants
           Proc. Natl. Acad. Sci. U.S.A. 97 (9), 4856-4861 (2000)
JOURNAL   20243785
MEDLINE   10781092
PUBMED    3 (bases 1 to 2898)
AUTHORS   Alfano,J.R. and Collmer,A.
TITLE     Direct Submission
SUBMITTED (07-FEB-2000) Dept. Biol. Sci., UNLV, 1854 Maryland
JOURNAL   Parkway, Las Vegas, NV 89154, USA

FEATURES             source
     1. .2898
     /organism="Pseudomonas syringae pv. syringae"
     /mol_type="genomic DNA"
     /strain="61"
     /db_xref="taxon:321"
     /notes="pathovar: syringae"
     256. .340
     /product="trna-Leu"
     708. .1046
     /note="ORF1"
     /codon_start=1
     /transl_table=11
     /product="putative type III chaperone"
     /protein_id="AAF71480.1"
     /db_xref="GI:8037768"
     /translations="MEMPALAFDDKGCACNMIDKAFALTLLRDDTHQRLLLIGLLEP
     HEDLPLORLLAGALNPVNAGPGIGHDEQSLGYHAYQSIPREKVSVMELKLEIAGLVE
     WMKCWEART"
     1106. .2233
     /gene="hrma"
     /notes="synonym: hopPsyA"
     1106. .2233
     /gene="hrma"
     /note="HopPsyA (Hrma)"
     /codon_start=1
     /transl_table=11
     /product="type III effector protein"
     /protein_id="AAF71481.1"
     /db_xref="GI:8037769"
     /translations="MNPPIHAFSSVEALRHSNVDIQAIKSEGOLEVNGKRYETRAAAD
     GSIAVLRPDQSKADKFFKGAARLIGGSQSQRAQIAQLVNEKAAAVPRLLDMLGRFDL
     EKGGSVAGAAIKAADSRLLTSKQTFAFQWAEKAEALGRYRNRYRLHDLQEGHARHNA
     YECGRVKNITWKYRLSITRKLTSLYAPQIHDDREEDLDLGRYIAEDRNARTGFRMV

```


REMARK
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES

Marburg, FRG
Revised by [3]
3 (bases 1 to 3719)
Vorholt, J.A.
Direct Submission
Submitted (15-JAN-1997) J.A. Vorholt, Max-Planck-Institut f.
terrestrische Mikrobiologie, Karl-von-Frisch-Strasse, D- 35043
Marburg, FRG
On Jan 16, 1997 this sequence version replaced gi:1419620.
Location/Qualifiers
1. .3719
/organism="Methanopyrus kandleri"
/mol_type="genomic DNA"
/strain="AV19/DSM 6324"
/db_xref="taxon:2320"
/clone_lib="lambda ZAP"
185. .940
/gene="orf6"
185. .940
/gene="orf6"
/codon_start=1
/transl_table=11
/protein_id="CAA67409.1"
/db_xref="GI:1781025"
/db_xref="SPTREMBL:P94940"
/translations="MRVLGVADGSHLKCAIVEDGSLDHTVVESTGTPVKVLRRLD
ELGAGIDFEDVTAVTCYGREALSDEFEVPELPAVALGASQLVEGARTVIDVGGQDT
KMKVEDGVDFQVNDKCAAGTGRFVENVCRRLGIEMSEVDEHAGADDPVKINSMC
AVFAETEVLNVRGIDVERILLGLVDSVAERVATMDIKVSPPEVVLVGMARCRVF
AELLSDRLEMGINVPNEAHVAGAFGAALWLEK"
1043. .2347
/gene="fwcB"
1043. .2347
/gene="fwcB"
/EC_number="1.2.99.5"
/codon_start=1
/transl_table=11
/product="beta-subunit of formylmethanofuran
dehydrogenase"
/protein_id="CAA67410.1"
/db_xref="GI:1781026"
/db_xref="SPTREMBL:P94941"
/translations="MGRKVKIDVVCFCGTLCDLLEVVVEDGEIVVEVRIACRIGAAKF
LTAQEDHRRTEPMKENGKWKIDYEDAAETARILVEAKWLMYGMSSSLCEAHVVG
IEIAKGVADINTASVCGPSTGLQDAGVPCSTGLGEVKNRADVIIYMGCPNMAHP
RHSRYTATFRGREDRTIIIVDPRTATABELADVHLQVRPNEDVELISALRA
AVHGIEIEEEVAGIPEVAEHVADLIKEASFGTLFGWGITWGRGRHRNIDNAICLI
RDLNEVAKWTLIMRGHYNVTGFNEVLTTTGYPAVDPSRGYPVYNGEFSADVLI
RGEVDAFVILASDPAHFPKAVEHMAWILPVCDPHWPTTAEIADLYVPVTLGIEW
EGTAYRMDSPVIMRKVPEPPESMLNDVLEWVIEKVBEMV"
complement (2355. .3620)
/gene="orf7"
complement (2355. .3620)
/gene="orf7"
/codon_start=1
/transl_table=11
/protein_id="CAA67411.1"
/db_xref="GI:1781027"
/db_xref="SPTREMBL:P94942"
/translations="MDIREIKELGIDILRERHDELLEALPEVVEEIIFLSOENRPMYS
FOEVADIHGARVELYEMREQKPIILTCFYVPEIIVLAAGVCVGLCGGAFFPI
DAEKLPRDLCLIKSFGFLVSRLCPCYQVATVIGENTCDGKKMYEIMSEHKDYY
VMEPMQDDEGLEWYHEQLIKFEVEELSNEITYESLLDAIERVNAKREAFRLY
ELRKHDPAPISGRDANLIAIAFYDDVDRFTKEVNELEKREVEEGEVAEAPRI
LVAGTQPMPIPHWKLIVYSCGVAVVEESCTCTRYFEREVSTEGDDVDLKNIAEA
YMETKCAIFTPNDVRKDIKKYKEWNCQGVILYLNKFCQPTAVESHKIESLRREGI
PALKLESDDSEEDVEQLKTRITFLESIA"
BASE COUNT 808 a 996 c 1060 g 855 t
ORIGIN

Query Match
Best Local Similarity 48.1%; Score 25; DB 1; Length 3719;
Matches 34; Conservative 69.4%; Pred. No. 2.le+03;
0; Mismatches 15; Indels 0; Gaps 0;

Qy
Db

RESULT 40
AE010324
LOCUS

DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL

FEATURES
source

gene

CDS

gene

CDS

gene

CDS

1 ATGCAGTTCTCTCAGCCTCTCATCGCTCTCGTCTCGCTCGCGCCTCCGCA 49
|||||
418 AGGAAGCTCGGCACGGCTCTCATCGAACTCGTCTCGCTCAGACCTCCCTA 370
|||||

AE010324 10264 bp DNA linear BCT 01-AUG-2002
Methanopyrus kandleri AV19 section 23 of 157 of the complete
genome

AE010324 AE009439
AE010324.1 GI:19886587

Methanopyrus kandleri AV19
Methanopyrus kandleri AV19

Archaea; Euryarchaeota; Methanopyri; Methanopyrales;
Methanopyraceae; Methanopyrus.

Slesarev,A.I., Mezheva,K.V., Makarova,K.S., Polushin,N.N.,
Shcherbinina,O.V., Shakhova,V.V., Belova,G.I., Aravind,L.,
Natale,D.A., Rogozin,I.B., Tatusov,R.L., Wolf,Y.I., Stetter,K.O.,
Malykh,A.G., Koonin,E.V. and Kozyavkin,S.A.
The Complete Genome of the Hyperthermophile Methanopyrus kandleri
AV19 and Monophyly of Archaeal Methanogens
Unpublished

2 (bases 1 to 10264)

Slesarev,A.I., Mezheva,K.V., Makarova,K.S., Polushin,N.N.,
Shcherbinina,O.V., Shakhova,V.V., Belova,G.I., Aravind,L.,
Natale,D.A., Rogozin,I.B., Tatusov,R.L., Wolf,Y.I., Stetter,K.O.,
Malykh,A.G., Koonin,E.V. and Kozyavkin,S.A.
Direct Submission
Submitted (04-FEB-2002) Fidelity Systems, Inc., Gaithersburg, MD
20879

Location/Qualifiers

1. .10264
/organism="Methanopyrus kandleri AV19"
/mol_type="genomic DNA"
/strain="AV19"
/db_xref="taxon:190192"
complement (82. .837)
/gene="MK0260"
complement (82. .837)
/gene="MK0260"
/codon_start=1
/transl_table=11
/product="Activator of 2-hydroxyglutaryl-CoA dehydratase,
contains a HSP70-class ATPase domain"
/protein_id="AAM01477.1"
/db_xref="GI:19886588"
/translations="MRVLGVADGSHLKCAIVEDGSLDHTVVESTGTPVKVLRRLD
ELGAGIDFEDVTAVTCYGREALSDEFEVPELPAVALGASQLVEGARTVIDVGGQDT
KMKVEDGVDFQVNDKCAAGTGRFVENVCRRLGIEMSEVDEHAGADDPVKINSMC
AVFAETEVLNVRGIDVERILLGLVDSVAERVATMDIKVSPPEVVLVGMARCRVF
AELLSDRLEMGINVPNEAHVAGAFGAALWLEK"
1028. .1489
/gene="MK0261"
1028. .1489
/gene="MK0261"
/codon_start=1
/transl_table=11
/product="Uncharacterized membrane protein, conserved in
archaea"
/protein_id="AAM01478.1"
/db_xref="GI:19886589"
/translations="MARVSPVLTVGLVAGVSAALQAVFKVPPPPAYGICIACTHRLV
NMIVNHAAGTTLGMAPVSPFVLTGVIGIFALIAFAHGFRIIRTRTHPIPLGVLG
FLVINPALFMGGCPIRTSLRTMYGDVALLGLISIAVGILGAFFYLKRKA"
1504. .1977
/gene="MK0262"
1504. .1977
/gene="MK0262"
/codon_start=1

Job time : 1954 secs
